

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 21:02:07 ; Search time 46 Seconds

(without alignments)  
848.324 Million cell updates/sec

Title: US-10-521-420-1

Perfect score: 2424

Sequence: 1 MKSILGLADTTTITTDL.....TVKIATVMSVSDTSAEAL 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5 COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/6 COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/H COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2407	99.3	472	2	US-09-826-509-469
2	1571.5	64.8	345	1	US-08-118-270-70
3	1571.5	64.8	345	4	PCT-US93-08528-70
4	705	29.1	360	2	US-09-826-509-471
5	365	15.1	393	4	PCT-US96-10618-3
6	364	15.0	364	2	US-09-169-205D-20
7	363	15.0	364	2	US-09-731-030A-15
8	363	15.0	364	2	US-09-811-838-2
9	360	14.9	364	4	PCT-US96-10618-2
10	355	14.6	364	2	US-08-763-938-2
11	353	14.6	381	1	US-08-845-566-3
12	353	14.6	381	1	US-08-467-948A-28
13	353	14.6	381	2	US-08-852-824-18
14	353	14.6	381	2	US-08-467-947A-28
15	353	14.6	381	2	US-09-731-030A-17
16	353	14.6	381	2	US-09-518-383-18
17	353	14.6	381	4	PCT-US96-10618-4
18	352.5	14.5	378	2	US-09-082-088-2
19	352.5	14.5	378	2	US-09-546-117-2
20	351.5	14.5	382	2	US-09-262-477-2
21	349	14.4	470	1	US-08-466-906B-8
22	349	14.4	470	2	US-09-201-746-8
23	346.5	14.3	378	2	US-09-169-205D-22
24	346.5	14.3	378	2	US-09-731-030A-18
25	346	14.3	382	2	US-09-542-733-2
26	344	14.2	382	2	US-09-169-205D-21
27	341	14.1	353	2	US-09-731-030A-16

Sequence 6, Appli	353	2	US-09-811-838-6
Sequence 7861, Ap	359	2	US-09-949-016-7861
Sequence 2, Appli	354	2	US-09-325-897-2
Sequence 4, Appli	354	2	US-09-837-726-2
Sequence 2, Appli	383	1	US-08-196-989B-4
Sequence 4, Appli	383	1	US-08-760-936-4
Sequence 14, Appli	383	2	US-09-225-024-4
Sequence 4, Appli	383	2	US-08-997-803-14
Sequence 11, Appli	384	2	US-08-852-824-4
Sequence 13, Appli	384	2	US-09-731-030A-11
Sequence 73, Appli	384	2	US-09-518-383-4
Sequence 73, Appli	384	2	US-09-731-030A-13
Sequence 73, Appli	334	4	US-08-118-270-73
Sequence 2, Appli	334	4	PCT-US93-08528-73
Sequence 2, Appli	351	2	US-08-861-747-2
Sequence 1, Appli	382	2	US-09-274-752D-1
Sequence 4, Appli	351	2	US-08-789-982-2
Sequence 4, Appli	351	2	US-09-811-838-4

ALIGNMENTS

RESULT 1

US-09-826-509-469  
; Sequence 469, Application US/09826509  
; Patent No. 6806054  
; GENERAL INFORMATION:  
; APPLICANT: Lehmann-Bruinsma, Karin  
; APPLICANT: Liauw, Chen W.  
; TITLE OF INVENTION: No. '6806054' Endogenous Constitutively Activated Known G  
; TITLE OF INVENTION: Protein-Coupled Receptors  
; FILE REFERENCE: AREN-207  
; CURRENT APPLICATION NUMBER: US/09/826,509  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: 60/195,747  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: 09/170,496  
; PRIOR FILING DATE: 1998-10-13  
; NUMBER OF SEQ ID NOS: 589  
; SOFTWARE: PatentIn Version 2.1  
; SEQ ID NO 469  
; LENGTH: 472  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-826-509-469

Query Match	99.3%	Score 2407;	DB 2;	Length 472;
Best Local Similarity	99.4%	Pred. No. 3.9e-194;		
Matches 469;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1	MKSILGLADTTTITTTD	LLYVGSNDIQYED	IKGDMASKLGYFPQKPLTSFRGSPQE 60
Db	1	MKSILGLADTTTITTTD	LLYVGSNDIQYED	IKGDMASKLGYFPQKPLTSFRGSPQE 60
Qy	61	KMTAGDNPQLPADQVNI	TEFYNKSLSFKENENI	OCGFNMFIECFNMFNPSQOLAIA 120
Db	61	KMTAGDNPQLPADQVNI	TEFYNKSLSFKENENI	OCGFNMFIECFNMFNPSQOLAIA 120
Qy	121	VLSITLGTFTVLENL	LVLCVILHSRSLRCR	PSYHFIGSLAVADLLGSVIFVYSFIDFHFV 180
Db	121	VLSITLGTFTVLENL	LVLCVILHSRSLRCR	PSYHFIGSLAVADLLGSVIFVYSFIDFHFV 180
Qy	181	HRKDSRNVPFLKGGV	TASFTASVGSGLFL	TAIAIYISIHRLAYKRIVTRPKAVAFCLM 240
Db	181	HRKDSRNVPFLKGGV	TASFTASVGSGLFL	TAIAIYISIHRLAYKRIVTRPKAVAFCLM 240
Qy	241	WTIAIVAVLPILGNCK	ELQSCVDIPPHIDET	YLMFWIGVTSVLLLPFIVYAVYIWLK 300
Db	241	WTIAIVAVLPILGNCK	ELQSCVDIPPHIDET	YLMFWIGVTSVLLLPFIVYAVYIWLK 300
Qy	301	AHSHAVRMIRGTQKSI	IIHTSEDKVQVTRPD	QARMAIRLAKTLVLVLIIICWGPPL 360

Db 301 AHSHAVRM1QRGTQKSIHHTSBDGKVQVTRPDQARMDIRLKKTLVLVLVLIICWGPLL 360  
Qy 361 AIMVYDFVFGKMKLIKTVPFACSMCLLNSTVNPPIIYALRSKDLRHAFSPSCGTAQ 420  
Db 361 AIMVYDFVFGKMKLIKTVPFACSMCLLNSTVNPPIIYALRSKDLRHAFSPSCGTAQ 420  
Qy 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
Db 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 2  
US-08-118-270-70  
; Sequence 70, Application US/08118270  
; Patent No. 5508384  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Randall B.  
; APPLICANT: Schuster, David I.  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/118,270  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 345 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-118-270-70

Query Match 64.8%; Score 1571.5; DB 1; Length 345;  
Best Local Similarity 90.8%; Pred. No. 4.1e-124;  
Matches 324; Conservative 8; Mismatches 12; Indels 13; Gaps 7;  
Qy 117 LAIAVLSLT-LGFTTVLENLLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGVSIFVYSFI 175  
Db 1 LAIAVLSLTLLGFTTVLENLLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGVSIFVYSFV 60  
Qy 176 DFHFVHRKDSRNVELFKLGGVTASTASVGSFLTAIARYISIHRLPLAYKRIIVTRPKAVV 235  
Db 61 DFHFVHRKDSRNVELFKLGGVTASTASVGSFLTAIDRYISIHPIATKRIIVRRPKAVV 120  
Qy 236 AFCIMMTIAIVIAVPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVYAYM 295  
Db 121 AFCIM-TIAIVIAVPLLGWNCCKLQSVCCDIFPLIDGTGLMFWIGVTSVLLLFIVYAYM 179

Qy 296 YILWKAHSHAVRM1QRGTQKSIHHTSBDGKVQVTRPDQARMAIRLAKTLVLVLVLIIC 355  
Db 180 YILWKAHSHAVR-AQGTQKSIHHTSBDGKVQVTRPDQARMDIRLAKTLVLVLVLIIC 238  
Qy 356 WGPLAIAMVYDFVFGKMKLIKTVPFACSMCLLNSTVNPPIIYALRSKDLRHAFSPMPSC 415  
Db 239 WGPLAIAMVYDFVFG---LLIKTVFAPCSU--LINSVNPPIIYALRSKDLRHAFRS-WPSC 292  
Qy 416 EGTAQPLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
Db 293 EGTAQPLDNSMGDSCLHKKHANNTASMHRAESCIKSTVKLA---LVSTDTSAEAL 345

RESULT 3  
PCT-US93-08528-70  
; Sequence 70, Application PC/TUS9308528  
; GENERAL INFORMATION:  
; APPLICANT: New York University  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/08528  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 345 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US93-08528-70

Query Match 64.8%; Score 1571.5; DB 4; Length 345;  
Best Local Similarity 90.8%; Pred. No. 4.1e-124;  
Matches 324; Conservative 8; Mismatches 12; Indels 13; Gaps 7;  
Qy 117 LAIAVLSLT-LGFTTVLENLLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGVSIFVYSFI 175  
Db 1 LAIAVLSLTLLGFTTVLENLLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGVSIFVYSFV 60  
Qy 176 DFHFVHRKDSRNVELFKLGGVTASTASVGSFLTAIARYISIHRLPLAYKRIIVTRPKAVV 235  
Db 61 DFHFVHRKDSRNVELFKLGGVTASTASVGSFLTAIDRYISIHPIATKRIIVRRPKAVV 120  
Qy 236 AFCIMMTIAIVIAVPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVYAYM 295  
Db 121 AFCIM-TIAIVIAVPLLGWNCCKLQSVCCDIFPLIDGTGLMFWIGVTSVLLLFIVYAYM 179  
Qy 296 YILWKAHSHAVRM1QRGTQKSIHHTSBDGKVQVTRPDQARMAIRLAKTLVLVLVLIIC 355

Db 180 YILKHAHSHAVR-AQRTQKSIHHTSEDKVQVTRPDQARMDIRLAKTLVLVLVLIIC 238  
Qy 356 WGPLLAIMVYDVGKMKLKTVPFACSMCLLNSTVNPPIYALRSKDLRHAFRSPSC 415  
Db 239 WGPLLAIMVYDVGKMKLKTVPFACSMCLLNSTVNPPIYALRSKDLRHAFRSPSC 292  
Qy 416 EGTAQPLDMSGSDCLHGHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
Db 293 EGTAQPLDMSGSDCLHGHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 345

## RESULT 4

US-09-826-509-471  
; Sequence 471, Application US/09826509  
; Patent No. 6806054  
; GENERAL INFORMATION:  
; APPLICANT: Lehmann-Bruinema, Karin  
; APPLICANT: Liaw, Chen W.  
; APPLICANT: Lin, I-Lin  
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G  
; TITLE OF INVENTION: Protein-Coupled Receptors  
; FILE REFERENCES: AREN-207  
; CURRENT APPLICATION NUMBER: US/09/826,509  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: 60/195,747  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: 09/170,496  
; PRIOR FILING DATE: 1998-10-13  
; NUMBER OF SEQ ID NOS: 589  
; SOFTWARE: PatentIn Version 2.1  
; SEQ ID NO 471  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-826-509-471

Query Match 29.1%; Score 705; DB 2; Length 360;  
Best Local Similarity 44.8%; Pred. No. 2.7e-51;  
Matches 148; Conservative 57; Mismatches 95; Indels 30; Gaps 6;  
Qy 78 ITEFYNKSLSPFKNEENIQCGENFMDIECPMVLNPSQOLAIAVLTLTGFTVLEMLLV 137  
Db 6 VTEIANGSKDGLDN-----PKDYMILSGQKTAVALVCLTLGLLSALENAV 54  
Qy 138 LCVTLHRSRLCRPSYHFIFGSLAVADLLGSVIFVYSFIDFHFVHRKDSRNVPFLKGGVT 197  
Db 55 LYLILSSHQLRRKPSYLFIFGSLAGADFLASVVFACSFVNFHVFHGVDSKAVFLKIGSVT 114  
Qy 198 ASFTASVGSLEPLTAIARYISHRPLAYKRIUTRPKANVAFCLMWTIAIAVIAVPLGWC 257  
Db 115 MTFTASVGSLLTAIDRYLCRLYPSPYKALITRGRALVTGLIMMVLGALVSLPLMGWTC 174  
Qy 258 EKLSQVSDIPPHIDETVLMFWIGVTSVLLLFIVVAYMILWKAHSHAVRMIOGTOKSI 317  
Db 175 --CRPCELPFLIPNDVLLSWLLFIAFLFGSIIYTYGHVWLKHAQVHA-----SL 223  
Qy 318 IIHTSEDKVQVTRPDQARMA--IRLAKTLVLVLVLIICWGPLLAIMVYDVGKMKLI 375  
Db 224 SGH--QDRQV---PGMARMLDVLKKTGLGLVLAVLICFWFVPLMALMAHSLATLSDQV 277  
Qy 376 KTVFACSMCLLNSTVNPPIYALRSKDLR 405  
Db 278 KKAFAFCSMLCLNSTMVNPVIALRSGBIR 307

## RESULT 5

PCT-US96-10618-3  
; Sequence 3, Application PC/TUS9610618  
; GENERAL INFORMATION:  
; APPLICANT: Coleman, Roger  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Au-Young, Janice

; APPLICANT: Bandman, Olga  
; APPLICANT: Sellhamer, Jeffrey J.  
; TITLE OF INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10618  
; FILING DATE: 20-JUN-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/000,352  
; FILING DATE: 20-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/567,817  
; FILING DATE: 06-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Glaister, Debra J.  
; REGISTRATION NUMBER: 33,888  
; REFERENCE/DOCKET NUMBER: PF-0042 PCT  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 393 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: U18405  
; PCT-US96-10618-3

Query Match 15.1%; Score 365; DB 4; Length 393;  
Best Local Similarity 27.2%; Pred. No. 1.1e-22;  
Matches 109; Conservative 71; Mismatches 176; Indels 44; Gaps 10;  
Qy 59 QEKMTAGDNQLPVADQVNITEFYFNKSLSPFKNEENIQCGENFMDIECPMVLNPSQOLA 118  
Db 13 QPQFTANPEQC-----FYNESIAFPYNSGKYLATE-----WNTVSKLV 52  
Qy 119 IAVLSLTGFTVLEMLLVLCVILHRSRLCRPSYHFIFGSLAVADLLGSVIFVYSFIDF 178  
Db 53 MG-LGITVCIFIMLANLVVAVIYVNRHFE-PIYILMANLAADDFAGLAYFYLMFNTG 110  
Qy 179 VFHRKDSRNVPFLKGGVTAFTASVGSLEPLTAIARYISHRPLAYKRIUTRPKANVAF 238  
Db 111 PNTRRLTVSTWLLRQGLIDTTVTASVANLLAIAIERHITVFRMQLHTRMSNR-RVVVVIV 169  
Qy 239 LMWTIAIVIAVPLGWCNCKLQSVCSDFIPPHIDETVLMFWIGVTSVLLLFIVVAYMIL 298  
Db 170 VITWMAIVMGAIPSVGNWNCIDENCSMAELYSDSLTVFWAIPLVTFVWVVLVLAHIF 229  
Qy 299 WKAHSHAVRMIOGTOKSIHHTSEDKVQVTRPDQARMAIRLAKTLVLVLVLIICWGP 358  
Db 230 GYVQRTWMSR-----HSSGPRENRT-----MWSLLKTVIVLGAFLICWTP 273  
Qy 359 LLAIMVYDVGKMKLKTVPFACSMCLLNSTVNPPIYALRSKDLRHAFRSPSC 416  
Db 274 GLVLLLDVCCPQCDVL-AYEKFFLLLAEFNSAMNPIIYSDYRDKEMSAFTRQILCCQSR 332



```
; ORGANISM: Homo sapiens
US-09-811-838-2

Query Match      15.0%; Score 363; DB 2; Length 364;
Best Local Similarity 27.2%; Pred. No. 1.5e-22;
Matches 108; Conservative 72; Mismatches 167; Indels 50; Gaps 10;

QY 59 QERKTAGDNPQVPADQVNITEFYKSLGSKFENEENIQCGENFMDIECFMVLNPSQOLA 118
DB 13 QPQFTANNEPQC-----FYNESIAFFYNRSGKHLATE-----WNTVSKLV 52

QY 119 IAVLSLTGLTFTVLENLVLCVILHRSRLCRPSYHFIGSLAVADLLGVSIFVYSFIDFH 178
DB 53 MG-LGITVCIFIMLANLVVVAIVNRRPHF-PIYVLMANLAAADPPAGLAYFYLMTGT 110

QY 179 VFHRKOSRVLPFKLGVTASTASVGSFLTAIARIYISIHRLPLAYKRIIVTRPKAVVAF 238
DB 111 PNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAIERHITVFRMLHTRMSNR-RVVVVIV 169

QY 239 LMWTIAIVIAVPLGLGNCKLOSCVDIPPHIDETVLMFWIGVTSVLLLFIVVAYMYIL 298
DB 170 VIWTAIVMGAIPSVGNWCIDENCNMAPLYSDSYLVFWAIFNLVTFVVMVLYAHIF 229

QY 299 WKAHSHAVRMIOQTOKSIIHTSECGKVQVTRPDQARMAIRLAKTILVLILVLIICWGP 358
DB 230 GYVQRTMRMR-----HSSGPRNRDT-----MMSLLKTVVIVLGFIICTWP 273

QY 359 LLAIMVYDFGKGNKLIKTVPAFCMCLLNSTVNPPIYALRSKDLRHAFRSMFPSCGT 418
DB 274 GLVLLLDVCCPQCDVL-AYEKFLLAEFNSAMNPIIYSYRDKEMSATFRQIL-----C 327

QY 419 AQPLDMSGSDCLHKNAN-----AASVRAAESCI 450
DB 328 QRSNPTEGTESSDRSASSLNHTILAGVHSNDHSVV 364

; RESULT 9
PCT-US96-10618-2
; Sequence 2, Application PC/TUS9610618
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Guegler, Karl J.
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10618
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,352
; FILING DATE: 20-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/567,817
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J.
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: PP-0042 PCT

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Rheumatoid Synovium
; CLONE: 80853
; PCT-US96-10618-2

Query Match      14.9%; Score 360; DB 4; Length 364;
Best Local Similarity 27.9%; Pred. No. 2.6e-22;
Matches 110; Conservative 70; Mismatches 170; Indels 44; Gaps 11;

QY 59 QERKTAGDNPQVPADQVNITEFYKSLGSKFENEENIQCGENFMDIECFMVLNPSQOLA 118
DB 13 QPQFTANNEPQC-----FYNESIAFFYNRSGKHLATE-----WNTVSKLV 52

QY 119 IAVLSLTGLTFTVLENLVLCVILHRSRLCRPSYHFIGSLAVADLLGVSIFVYSFIDFH 178
DB 53 MG-LGITVCIFIMLANLVVVAIVNRRPHF-PIYVLMANLAAADPPAGLAYFYLMTGT 110

QY 179 VFHRKOSRVLPFKLGVTASTASVGSFLTAIARIYISIHRLPLAYKRIIVTRPKAVVAF 238
DB 111 PNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAIERHITVFRMLHTRMSNR-RVVVVIV 169

QY 239 LMWTIAIVIAVPLGLGNCKLOSCVDIPPHIDETVLMFWIGVTSVLLLFIVVAYMYIL 298
DB 170 VIWTAIVMGAIPSVGNWCIDENCNMAPLYSDSYLVFWAIFNLVTFVVMVLYAHIF 229

QY 299 WKAHSHAVRMIOQTOKSIIHTSECGKVQVTRPDQARMAIRLAKTILVLILVLIICWGP 358
DB 230 GYVQRTMRMR-----HSSGPRNRDT-----MMSLLKTVVIVLGFIICTWP 273

QY 359 LLAIMVYDFGKGNKLIKTVPAFCMCLLNSTVNPPIYALRSKDLRHAFRSMFPSCGT 418
DB 274 GLVLLLDVCCPQCDVL-AYEKFLLAEFNSAMNPIIYSYRDKEMSATFRQIL-----CORS 331

; RESULT 10
US-08-763-938-2
; Sequence 2, Application US/08763938
; Patent No. 6140060
; GENERAL INFORMATION:
; APPLICANT: CHUN, Jerold J.M.
; APPLICANT: HECHT, Jonathan H.
; TITLE OF INVENTION: CLONED LYSOPHOSPHATIDIC ACID
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray and Oram LLP
; STREET: 655 15th Street, N.W., Suite 330 - G St. Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/763,938
```



```
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,948A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04079
FILING DATE: 30-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-467-948A-28

Query Match 14.6%; Score 353; DB 1; Length 381;
Best Local Similarity 27.3%; Pred. No. 1.1e-21;
Matches 102; Conservative 71; Mismatches 129; Indels 72; Gaps 12;

QY 68 POLVP-----ADQVN---ITEFVN---KSLSSFKENEENIQGFNMDIECFMVL 111
DB 3 PTVPLVKHRSSVSDVYNDIIVRHNYNKGKLNISADKENSIKL-----47
QY 112 NPSQQLAIAVLSLTGLTFTVLENLVLCVILHSRLCRPSYHFGISLAVADLLGSVIFV 171
DB 48 -----TSVVFILICCFIILENIFVLLTIWTKKFH-RPMYFIGNLSLSDLLAGVAYT 99
QY 172 YSFIDFVHFRKDSRNVPFLKGGVTASFTASVGSFLTAIARYIS-----IHRPLAYKR 226
DB 100 ANLLSGATTYKLTTPAQWFLREGSMFVALSASFSLAIAIERYITMLKMKLHNGSNFR 159
QY 227 IVTRPKAVAFCLMWTIAIVLPLLGWCKEKLQSCDIFPHIDETILMFVIGVTSVL 286
DB 160 LP----LLISAC--WVLSILGLPIMGWNCISALSSCSVLPDYHKHYILFCTTVFTLL 213
QY 287 LLFTV--YAYWYILKASHAVRMIOQTOKSIIHTSDEGKVQVTRPDQARMAIRLAKT 344
DB 214 LLSIVLYCRYSLVTRSRRLTF-----RKNISKASRSSENVALLKT 256
QY 345 LVLIIVLIIICWGPLLAIMVYDFGKMKLIKTVF--AFCSMLCLLNSTVNPPIIYALRSK 402
DB 257 VIIVLSVFIACWAPLFIILLDDVGCKV-KTCDILFRAEYFLVLAVLSNGTNPPIIYTLTK 315
QY 403 DLRAFRSMFSPCE 416
DB 316 EMRRAFIRMSCK 329

RESULT 14
US-08-467-947A-28
Sequence 28, Application US/08467947A
Patent No. 6090575
GENERAL INFORMATION:
APPLICANT: LI, YI
APPLICANT: CAO, LIANG
APPLICANT: NI, JIAN
APPLICANT: GENTZ, REINER
APPLICANT: SUTTON III, GRANGER G.
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
TITLE OF INVENTION: Coupled Receptor GPR1
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,947A
```

/ FILING DATE: 06-JUN-1995  
/ CLASSIFICATION: 435  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/US95/04079  
/ FILING DATE: 30-MAR-1995  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: STEFFE, ERIC K.  
/ REGISTRATION NUMBER: 36,698  
/ REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 202-371-2600  
/ TELEFAX: 202-371-2540  
/ INFORMATION FOR SEQ ID NO: 28:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 381 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: not relevant  
/ MOLECULE TYPE: peptide  
/ US-08-467-947A-28

Query Match 14.6%; Score 353; DB 2; Length 381;  
Best Local Similarity 27.3%; Pred. No. 1.1e-21;  
Matches 102; Conservative 71; Mismatches 129; Indels 72; Gaps 12;

QY 68 PQLVP-----ADQVN---ITEFYN---KSLSSFKENEENIQCGENFMDIECFMVL 111  
DB 3 PTVPLVKAHRSSVDYVNDIIVRHNYVTGKLNISADKENSIKL-----47  
QY 112 NPSQQLAIAVLSLTGFTFTVLENLLVLCVILHSRLCRPSYHFHIGSLAVADLLGSVIFV 171  
DB 48 -----TSVVFILICCFIENIFVLLTIWTKKPH-RPMYFYFIGNLALSOLLAGVAYT 99  
QY 172 YSFIDFVHVRKDSRVNLFKLGVTASTASVGSILFLTAIARIYS-----IHRPLAYKR 226  
DB 100 ANLLLSGATTYKLTTPAQWFLREGSMFVALSASVFSLLAIAIERIITMLKMKLHNGSNFR 159  
QY 227 IVTRPKAVVAFCLMWTIAIVLPLLGWNCCKLQSVCSDFPHIDETVLMFHWIGVTSVL 286  
DB 160 LF-----LLISAC--WVLSILGLPLMGWNCISALSSCSTVPLVYHKHYILFCTTVFTLL 213  
QY 287 LFLFIV--YAYMYILWKAHSHAVRMIOQTQKSIHHTSBDGKVQVTRPDQARMALRAKT 344  
DB 214 LLSIVILYCRISLVTRSRRLTF-----RKNISKASRSSSENVALLKT 256  
QY 345 LVLLLVLLIICWGPLLAIMVYDFGKMNKLIKTVF--AFCSMLCLLNSTVNPILIIYALRSK 402  
DB 257 VIIVLSVFIACWAPLFIILLDDVGCKV-KTCDILFRAEYFLVLAVLNSGTNPILIIYTLTK 315  
QY 403 DLRHAFRSMFPSC 416  
DB 316 EMRRAFIRMSCK 329

RESULT 15  
US-09-731-030A-17  
/ Sequence 17, Application US/09731030A  
/ Patent No. 6566096  
/ GENERAL INFORMATION:  
/ APPLICANT: MUNROE, Donald G  
/ APPLICANT: GUPTA, Ashwani K.  
/ APPLICANT: ZASTAWNY, Roman L.  
/ TITLE OF INVENTION: MAMMALIAN EDG-7 RECEPTOR HOMOLOGS  
/ FILE REFERENCE: 8074-0015  
/ CURRENT APPLICATION NUMBER: US/09/731,030A  
/ CURRENT FILING DATE: 1998-12-29  
/ PRIOR APPLICATION NUMBER: 60/070,184  
/ PRIOR FILING DATE: 1997-12-30  
/ NUMBER OF SEQ ID NOS: 21  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 17  
/ LENGTH: 381  
/ TYPE: PRT

/ ORGANISM: Unknown  
/ FEATURE:  
/ OTHER INFORMATION: Description of Unknown Organism: Mammalian EDG  
/ US-09-731-030A-17  
Query Match 14.6%; Score 353; DB 2; Length 381;  
Best Local Similarity 27.3%; Pred. No. 1.1e-21;  
Matches 102; Conservative 71; Mismatches 129; Indels 72; Gaps 12;  
QY 68 PQLVP-----ADQVN---ITEFYN---KSLSSFKENEENIQCGENFMDIECFMVL 111  
DB 3 PTVPLVKAHRSSVDYVNDIIVRHNYVTGKLNISADKENSIKL-----47  
QY 112 NPSQQLAIAVLSLTGFTFTVLENLLVLCVILHSRLCRPSYHFHIGSLAVADLLGSVIFV 171  
DB 48 -----TSVVFILICCFIENIFVLLTIWTKKPH-RPMYFYFIGNLALSOLLAGVAYT 99  
QY 172 YSFIDFVHVRKDSRVNLFKLGVTASTASVGSILFLTAIARIYS-----IHRPLAYKR 226  
DB 100 ANLLLSGATTYKLTTPAQWFLREGSMFVALSASVFSLLAIAIERIITMLKMKLHNGSNFR 159  
QY 227 IVTRPKAVVAFCLMWTIAIVLPLLGWNCCKLQSVCSDFPHIDETVLMFHWIGVTSVL 286  
DB 160 LF-----LLISAC--WVLSILGLPLMGWNCISALSSCSTVPLVYHKHYILFCTTVFTLL 213  
QY 287 LFLFIV--YAYMYILWKAHSHAVRMIOQTQKSIHHTSBDGKVQVTRPDQARMALRAKT 344  
DB 214 LLSIVILYCRISLVTRSRRLTF-----RKNISKASRSSSENVALLKT 256  
QY 345 LVLLLVLLIICWGPLLAIMVYDFGKMNKLIKTVF--AFCSMLCLLNSTVNPILIIYALRSK 402  
DB 257 VIIVLSVFIACWAPLFIILLDDVGCKV-KTCDILFRAEYFLVLAVLNSGTNPILIIYTLTK 315  
QY 403 DLRHAFRSMFPSC 416  
DB 316 EMRRAFIRMSCK 329

Search completed: January 6, 2006, 21:13:17  
Job time : 48 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 21:00:32 ; Search time 189 seconds

(without alignments)  
1097.285 Million cell updates/sec

Title: US-10-521-420-1

Perfect score: 2424

Sequence: 1 MKSILDGLADTTFTTTDL.....TVKIAKVTMSVSDTSAEAL 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 21:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1980s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2000s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*  
9: Geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	472	ADH77048	Adh77048 hCB-1 D21
2	2412	99.5	472	AAR14196	Aar14196 Human can
3	2412	99.5	472	ABR58525	ABR58525 Human can
4	2412	99.5	472	ABP81800	ABP81800 Human can
5	2412	99.5	472	ADD46383	Add46383 Human Pro
6	2412	99.5	472	ADL15025	Adl15025 Human CB1
7	2412	99.5	472	ADN38918	Adn38918 Cancer/an
8	2412	99.5	472	ABO60051	AbO60051 Human Gen
9	2412	99.5	472	ADO29261	Ado29261 Human GPC
10	2412	99.5	472	AEA81221	Aea81221 Human can
11	2407	99.3	472	ABB56338	Abb56338 Non-endog
12	2401	99.1	472	ADN38920	Adn38920 Cancer/an
13	2343.5	96.7	473	AAR14195	Aar14195 Rat canna
14	2343.5	96.7	473	ADD46381	Add46381 Rat Prote
15	2340.5	96.6	473	ABB57048	Abb57048 Mouse lsc
16	2340.5	96.6	473	ADO29262	Ado29262 Mouse GPC
17	2209.5	91.2	439	ADM82794	Adm82794 Human can
18	2132.5	88.0	789	ADU82881	Adu82881 Ligand up
19	1954	80.6	411	ADJ70504	Adj70504 Human bea
20	1954	80.6	411	ADN38922	Adn38922 Cancer/an
21	1954	80.6	411	AEA81222	Aea81222 Human can
22	1571.5	64.8	345	AAR48749	Aar48749 G-protein
23	1571.5	64.8	345	AAW02721	Aaw02721 G-protein
24	735.5	30.3	363	ADQ97953	Adq97953 Mouse can

## ALIGNMENTS

### RESULT 1

ADH77048

ID ADH77048 standard; protein; 472 AA.

XX AC ADH77048;

XX AC ADH77048;

DT 22-APR-2004 (first entry)

XX hCB-1 D213A double mutant, SEQ ID 1.

XX Anorectic; neuroprotective; cardiovascular; respiratory;  
gastrointestinal; cannabinoid; CB; receptor; obesity; psychiatric;  
neurological; immune; cardiovascular; reproductive; endocrine; disorder;  
respiratory; gastrointestinal; hCB-1 D213A; mutein.

OS Homo sapiens.

OS Synthetic.

XX WO2004008150-A1.

XX 22-JAN-2004.

PD 14-JUL-2003; 2003WO-GB003066.

PF 17-JUL-2002; 2002SE-00002242.

XX (ASTR ) ASTRAZENECA AB.

PA (ASTR ) ASTRAZENECA UK LTD.

XX Greasley P;

XX WPI; 2004-143121/14.

DR Identifying an inverse agonist of a cannabinoid (CB) receptor, useful in  
treating obesity, psychiatric and neurological disorders, comprises  
contacting a test inhibitory agent with constitutively active CB  
receptor.

XX Claim 31; SEQ ID NO 1; 31pp; English.

PS The invention relates to a method for identifying an inverse agonist of a  
cannabinoid (CB) receptor comprising contacting a CB receptor test  
inhibitory agent with the cell expressing the constitutively active CB  
receptor. The method is useful in identifying an inverse agonist of a CB  
receptor. The true antagonist or inverse agonist is useful in preparing a  
medicament for treating or preventing a disorder associated with a CB  
receptor. The disorder is obesity, psychiatric and neurological

Ado29264 Mouse GPC  
Aar67001 Cannabino  
Abp81801 Human can  
Abu04877 Human exp  
Abu04878 Human exp  
Abu04876 Human exp  
Abu04875 Human exp  
Ado29263 Human GPC  
Adq97956 Human can  
Abb56339 Non-endog  
Aaw03642 Human can  
Aar67002 Cannabino  
Abp54018 Human cen  
Ado28758 Human cen  
Ado05262 Central c  
Adx44577 Human cla  
Ady83813 Central c  
Aea13783 Human cen  
Abu67237 G-protein  
Abg75679 Human Edg  
Ado28787 Arrestin

25 734.5 30.3 347 8 ADO29264  
26 710 29.3 360 2 AAR67001  
27 710 29.3 360 6 ABP81801  
28 710 29.3 360 6 ABU04877  
29 710 29.3 360 6 ABU04878  
30 710 29.3 360 6 ABU04876  
31 710 29.3 360 6 ABU04875  
32 710 29.3 360 8 ADO29263  
33 709 29.2 360 8 ADQ97956  
34 705 29.1 360 4 ABB56339  
35 435 17.9 116 2 AAW03642  
36 413.5 17.1 180 2 AAR67002  
37 410 16.9 80 6 ABP54018  
38 410 16.9 80 8 ADO28758  
39 410 16.9 80 8 ADO05262  
40 410 16.9 80 9 ADX44577  
41 410 16.9 80 9 ADY83813  
42 410 16.9 80 9 AEA13783  
43 404 16.7 80 5 ABU67237  
44 369.5 15.2 362 5 ABG75679  
45 369.5 15.2 362 8 ADO28787

CC disorders. They are also useful in treating immune cardiovascular,  
 CC reproductive and endocrine disorders and also diseases related to  
 CC respiratory and gastrointestinal systems. The current sequence represents  
 CC the amino acid sequence of the hCB-1 D213A double mutant.

XX  
 SQ Sequence 472 AA;

Query Match 100.0%; Score 2424; DB 8; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-241; Indels 0; Gaps 0;  
 Matches 472; Conservative 0; Mismatches 0;

QY 1 MKSILDLGLADTTFRITITDLYVGSNDIQYEDIKGDMASKLGYFPQKPLTFRGSPFQE 60  
 DB 1 MKSILDLGLADTTFRITITDLYVGSNDIQYEDIKGDMASKLGYFPQKPLTFRGSPFQE 60

QY 61 KMTAGNPOLVPADQVNITEFYNKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 120  
 DB 61 KMTAGNPOLVPADQVNITEFYNKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 120

QY 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180  
 DB 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180

QY 181 HRKDSRVNLFKLGCVTASFTASVGSLEFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 DB 181 HRKDSRVNLFKLGCVTASFTASVGSLEFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240

QY 241 WTIAIVAVPLLGWNCCKLQSVCSDFPHIDTYLMFWIGTVSVLLLFIVYAYMYLWK 300  
 DB 241 WTIAIVAVPLLGWNCCKLQSVCSDFPHIDTYLMFWIGTVSVLLLFIVYAYMYLWK 300

QY 301 AHSHAVRMIQRTGQKSIITHTSDEGKVQVTRPDQARMAIRLAKTLVLILVLLIICWGPLL 360  
 DB 301 AHSHAVRMIQRTGQKSIITHTSDEGKVQVTRPDQARMAIRLAKTLVLILVLLIICWGPLL 360

QY 361 AIMVYDVFQGMNKLKTVFAFCMLCLLNSTVNPITIALRSKDLRHAFRSMPPSCBGTQA 420  
 DB 361 AIMVYDVFQGMNKLKTVFAFCMLCLLNSTVNPITIALRSKDLRHAFRSMPPSCBGTQA 420

QY 421 PLDMSGDSDDLKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 DB 421 PLDMSGDSDDLKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 2

AAR14196  
 ID AAR14196 standard; protein; 472 AA.

XX  
 AC AAR14196;

XX  
 DT 25-MAR-2003 (revised)

DT 17-DEC-2001 (revised)

DT 19-DEC-1991 (first entry)

XX  
 DE Human cannabinoid receptor.

XX  
 KW Cannabis sativa; marijuana; drug test; substance K receptor.

XX  
 OS Homo sapiens.

XX  
 PN USN7564075-N.

XX  
 PD 03-SEP-1991.

XX  
 PF 08-AUG-1990; 90US-00564075.

XX  
 PR 08-AUG-1990; 90US-00564075.

XX  
 PA (USSH ) NAT INST OF HEALTH.

XX  
 PA (USDC ) US SEC OF COMMERCE.

XX  
 PI Matsuda L, Brownstein M;

XX

DR WPI; 1991-303326/41.  
 DR N-PSDB; AAQ14003.

XX  
 PT DNA encoding mammalian cannabinoid receptor - used for producing receptor  
 for screening drugs and ligands and in detection.

XX  
 PS Disclosure; Fig 5; 25pp; English.

XX  
 CC SKR6 cDNA encoding the rat cannabinoid receptor was isolated from a rat  
 cerebral cortex cDNA library. An EcoRI-XbaI fragment was used to screen a  
 human cosmid library. A positive clone was identified and sequenced.  
 CC There is ca. 97 per cent homology between this amino acid sequence  
 deduced from the sequence of the positive human clone and rat cannabinoid  
 CC receptor (see AAR14195). Recombinantly produced receptor can be used to  
 CC screen for new drugs suitable for treatment of cannabinoid-treatable  
 CC conditions, e.g. glaucoma, bronchial asthma, etc. (Note: Revised entry  
 CC submitted to correct the patent number format of US Government-owned NDIS  
 CC applications to prevent clashes with ongoing US granted patent numbers.  
 CC For further information please visit the Derwent web site at  
 CC www.derwent.com/dwpi/updates/ntis\_us.html.) (Updated on 25-MAR-2003 to  
 CC correct PA field.)

XX  
 SQ Sequence 472 AA;

Query Match 99.5%; Score 2412; DB 2; Length 472;  
 Best Local Similarity 99.6%; Pred. No. 3.8e-240;  
 Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKSILDLGLADTTFRITITDLYVGSNDIQYEDIKGDMASKLGYFPQKPLTFRGSPFQE 60  
 DB 1 MKSILDLGLADTTFRITITDLYVGSNDIQYEDIKGDMASKLGYFPQKPLTFRGSPFQE 60

QY 61 KMTAGNPOLVPADQVNITEFYNKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 120  
 DB 61 KMTAGNPOLVPADQVNITEFYNKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 120

QY 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180  
 DB 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180

QY 181 HRKDSRVNLFKLGCVTASFTASVGSLEFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 DB 181 HRKDSRVNLFKLGCVTASFTASVGSLEFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240

QY 241 WTIAIVAVPLLGWNCCKLQSVCSDFPHIDTYLMFWIGTVSVLLLFIVYAYMYLWK 300  
 DB 241 WTIAIVAVPLLGWNCCKLQSVCSDFPHIDTYLMFWIGTVSVLLLFIVYAYMYLWK 300

QY 301 AHSHAVRMIQRTGQKSIITHTSDEGKVQVTRPDQARMAIRLAKTLVLILVLLIICWGPLL 360  
 DB 301 AHSHAVRMIQRTGQKSIITHTSDEGKVQVTRPDQARMAIRLAKTLVLILVLLIICWGPLL 360

QY 361 AIMVYDVFQGMNKLKTVFAFCMLCLLNSTVNPITIALRSKDLRHAFRSMPPSCBGTQA 420  
 DB 361 AIMVYDVFQGMNKLKTVFAFCMLCLLNSTVNPITIALRSKDLRHAFRSMPPSCBGTQA 420

QY 421 PLDMSGDSDDLKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 DB 421 PLDMSGDSDDLKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 3

ABR58525

ID ABR58525 standard; protein; 472 AA.

XX  
 AC ABR58525;

XX  
 DT 09-JUL-2003 (first entry)

XX  
 DE Human cannabinoid receptor 1 (brain) protein.

XX  
 KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;  
 heart disease; atherosclerosis; endometriosis.

XX OS Homo sapiens.  
XX PN WO2003025138-A2.  
XX PD 27-MAR-2003.  
XX PF 17-SEP-2002; 2002WO-US029560.  
XX PR 17-SEP-2001; 2001US-0323469P.  
XX PR 20-SEP-2001; 2001US-0323887P.  
XX PR 13-NOV-2001; 2001US-0350666P.  
XX PR 08-FEB-2002; 2002US-0355145P.  
XX PR 08-FEB-2002; 2002US-0355257P.  
XX PR 12-APR-2002; 2002US-0372246P.  
XX PA (BOSB-) EOS BIOTECHNOLOGY INC.  
XX PI Afar D, Aziz N, Gish KC, Hevesi PA, Mack DH, Wilson KE;  
PI Zlotnik A;  
XX WPI; 2003-354600/33.  
XX N-PSDB; ACC72645.  
XX New genes that are up-regulated or down-regulated in cancers, useful as  
XX markers for diagnosing e.g. cancer, ischemia or heart diseases, or as  
XX therapeutic targets for screening drugs for treating these diseases.  
XX Claim 12; Page 143; 767pp; English.  
XX The present invention describes an isolated nucleic acid molecule, which  
XX comprises the sequence of any of the genes that are up-regulated or down-  
XX regulated in specific cancers (e.g. about 1031 genes up-regulated in  
XX acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer  
XX related gene nucleotide sequences which encode the proteins given in  
XX ABR58521 to ABR58709. Also described: (1) determining the presence or  
XX absence of a pathological cell in a patient; (2) an expression vector  
XX comprising a nucleic acid molecule described above; (3) a host cell  
XX comprising the vector; (4) an isolated polypeptide, which is encoded by  
XX the nucleic acid; (5) an antibody that specifically binds the polypeptide  
XX of (4); (6) specifically targeting a compound to a pathological cell in a  
XX patient by administering to the patient the antibody above; and (7) a  
XX drug screening assay. The nucleic acid is useful as diagnostic markers or  
XX therapeutic targets. In particular, the nucleic acid is useful for  
XX diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,  
XX bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
XX pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
XX atherosclerosis and endometriosis. The nucleic acid is also useful in  
XX drug screening, particularly for identifying agents for treating these  
XX pathologies  
XX SQ Sequence 472 AA;  
Query Match 99.5%; Score 2412; DB 6; Length 472;  
Best Local Similarity 99.6%; Pred. No. 3.8e-240;  
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MKSILDGLADTTFRTITDLYVGSNDIQYEDIKDMASKLGYFPQKPLTSFRGSPQE 60  
Db 1 MKSILDGLADTTFRTITDLYVGSNDIQYEDIKDMASKLGYFPQKPLTSFRGSPQE 60  
Qy 61 KMTAGDNPOLPADQVNITEFYNKSLSFKEENIQCENFMDECFMVLNPSQOLAIA 120  
Db 61 KMTAGDNPOLPADQVNITEFYNKSLSFKEENIQCENFMDECFMVLNPSQOLAIA 120  
Qy 121 VLSLTGTFVLENLVLCVLHSLRSLCRPSYHPIGSLAVADLGSVIFYVSFDFHVF 180  
Db 121 VLSLTGTFVLENLVLCVLHSLRSLCRPSYHPIGSLAVADLGSVIFYVSFDFHVF 180  
Qy 181 HKDSRNVPFLKLGVTASVGSLSFLTAIARYISIHRLPLAYKRIIVTRPKAVVAFCLM 240  
Db 181 HKDSRNVPFLKLGVTASVGSLSFLTAIARYISIHRLPLAYKRIIVTRPKAVVAFCLM 240

Qy 241 WTIAIVIAVPLLGWNCCKLQSVCSDFPHIDETYLMTFMTGVTSLVLLLFVIYAYMYLWK 300  
Db 241 WTIAIVIAVPLLGWNCCKLQSVCSDFPHIDETYLMTFMTGVTSLVLLLFVIYAYMYLWK 300  
Qy 301 AHSHAVRMIOQTGKSIITHTSEDGKVQVTRPQARMAIRLAKTLVLVLLVLIICWGP 360  
Db 301 AHSHAVRMIOQTGKSIITHTSEDGKVQVTRPQARMAIRLAKTLVLVLLVLIICWGP 360  
Qy 361 AINVYDVFGWKNKLIKTVFAFCSCMLCLLNSTVNPPIIYALRSKDLRHAFRSMFPSC 420  
Db 361 AINVYDVFGWKNKLIKTVFAFCSCMLCLLNSTVNPPIIYALRSKDLRHAFRSMFPSC 420  
Qy 421 PLDSMGDSCLLHKHANNAAASVHRAAESCIKSTVKIAKVTMSVSTDTSAAAL 472  
Db 421 PLDSMGDSCLLHKHANNAAASVHRAAESCIKSTVKIAKVTMSVSTDTSAAAL 472  
RESULT 4  
ABP81800  
ID ABP81800 standard; protein; 472 AA.  
XX AC ABP81800;  
XX AC ABP81800;  
XX DT 04-MAR-2003 (first entry)  
XX DE Human cannabinoid receptor 1 protein SEQ ID NO:84.  
XX KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;  
KW G protein-coupled receptor modulator; antibody; immune-related disease;  
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;  
KW immunological-related cell proliferative disease; autoimmune disease;  
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;  
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;  
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;  
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;  
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;  
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;  
KW ulcer.  
XX OS Homo sapiens.  
XX WO200261087-A2.  
XX 08-AUG-2002.  
XX 19-DEC-2001; 2001WO-US050107.  
XX 19-DEC-2000; 2000US-0257144P.  
XX (LIFE-) LIFESPAN BIOSCIENCES INC.  
XX Burmer GC, Roush CL, Brown JP;  
XX WPI; 2003-046718/04.  
XX N-PSDB; ABZ42646.  
XX New isolated antigenic peptides e.g., for G protein-coupled receptors  
XX (GPCR), useful for diagnosing and designing drugs for treating conditions  
XX in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or  
XX autoimmune diseases.  
XX Disclosure; Fig 1; 523pp; English.  
XX The present invention describes antigenic peptides (I) comprising: (a)  
XX any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino  
XX acids. Also described: (1) an assay for the detection of a particular G  
XX protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;  
XX and (2) an isolated antibody having high specificity and high affinity or  
XX avidity for a particular GPCR. (1) can be used as GPCR modulators and in  
XX gene therapy. The antigenic peptides for GPCRs are useful in detecting an  
XX antibody against a particular GPCR, and in the production of specific  
XX antibodies. The peptides and antibodies are also useful for detecting the  
XX presence or absence of corresponding GPCRs. The antigenic peptides for

CC GPCRs and antibodies are useful for diagnosing and designing drugs for  
CC treating immune-related diseases, growth-related diseases, cell  
CC regeneration-related disease, immunological-related cell proliferative  
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,  
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,  
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute  
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host  
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,  
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory  
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, or  
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or  
CC any other disorder in which GPCRs are involved. The antibodies may be  
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode  
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the  
CC exemplification of the present invention  
XX  
SQ Sequence 472 AA;  
Query Match 99.5%; Score 2412; DB 6; Length 472;  
Best Local Similarity 99.6%; Pred. No. 3.8e-240;  
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MKSILDGLADTTFRITTTDLLLVGSDNIQYEDIKGMASKLGYFPQKFLTSFRGSPFQ 60  
DB |||||||  
QY 1 MKSILDGLADTTFRITTTDLLLVGSDNIQYEDIKGMASKLGYFPQKFLTSFRGSPFQ 60  
DB |||||||  
QY 61 KMTAGNPOLVPADQVNITEFYNNKLSLSPKNEENIQCGENFMDIECFVNLPSQOLA 120  
DB |||||||  
QY 61 KMTAGNPOLVPADQVNITEFYNNKLSLSPKNEENIQCGENFMDIECFVNLPSQOLA 120  
DB |||||||  
QY 121 VLSLTGTTFTVLENLVLCVILHSRLCRPSVHFITGSLAVADLLGSLVFVYSFIDFHV 180  
DB |||||||  
QY 121 VLSLTGTTFTVLENLVLCVILHSRLCRPSVHFITGSLAVADLLGSLVFVYSFIDFHV 180  
DB |||||||  
QY 181 HRKDSRNVELFKLGGVTSFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
DB |||||||  
QY 181 HRKDSRNVELFKLGGVTSFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
DB |||||||  
QY 241 WTIAIVAVPLLGWNCCKLQSCDIFPHIDETVLMFWIGVTSVLLFIVVAYMILWK 300  
DB |||||||  
QY 241 WTIAIVAVPLLGWNCCKLQSCDIFPHIDETVLMFWIGVTSVLLFIVVAYMILWK 300  
DB |||||||  
QY 301 AHSHAVRMIQRGTKSIITHTSDGKVVQVTRPDQARMALAKTLVLILVLIICWGP 360  
DB |||||||  
QY 301 AHSHAVRMIQRGTKSIITHTSDGKVVQVTRPDQARMALAKTLVLILVLIICWGP 360  
DB |||||||  
QY 361 AIMVYDVFQGMNKLITKTVAFCSMLCLNSTVNPITVALRSKDLRHAFRSMFPCB 420  
DB |||||||  
QY 361 AIMVYDVFQGMNKLITKTVAFCSMLCLNSTVNPITVALRSKDLRHAFRSMFPCB 420  
DB |||||||  
QY 421 PLDNSMGDSGDCILKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
DB |||||||  
QY 421 PLDNSMGDSGDCILKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
DB |||||||  
RESULT 5  
ADD46383  
ID ADD46383 standard; protein; 472 AA.  
XX  
AC ADD46383;  
XX  
XX 29-JAN-2004 (first entry)  
XX  
DE Human Protein P21554, SEQ ID NO 12063.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX

PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
DR GENBANK; P21554.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 472 AA;  
Query Match 99.5%; Score 2412; DB 7; Length 472;  
Best Local Similarity 99.6%; Pred. No. 3.8e-240;  
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MKSILDGLADTTFRITTTDLLLVGSDNIQYEDIKGMASKLGYFPQKFLTSFRGSPFQ 60  
DB |||||||  
QY 1 MKSILDGLADTTFRITTTDLLLVGSDNIQYEDIKGMASKLGYFPQKFLTSFRGSPFQ 60  
DB |||||||  
QY 61 KMTAGNPOLVPADQVNITEFYNNKLSLSPKNEENIQCGENFMDIECFVNLPSQOLA 120  
DB |||||||  
QY 61 KMTAGNPOLVPADQVNITEFYNNKLSLSPKNEENIQCGENFMDIECFVNLPSQOLA 120  
DB |||||||  
QY 121 VLSLTGTTFTVLENLVLCVILHSRLCRPSVHFITGSLAVADLLGSLVFVYSFIDFHV 180  
DB |||||||  
QY 121 VLSLTGTTFTVLENLVLCVILHSRLCRPSVHFITGSLAVADLLGSLVFVYSFIDFHV 180  
DB |||||||  
QY 181 HRKDSRNVELFKLGGVTSFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
DB |||||||  
QY 181 HRKDSRNVELFKLGGVTSFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
DB |||||||  
QY 241 WTIAIVAVPLLGWNCCKLQSCDIFPHIDETVLMFWIGVTSVLLFIVVAYMILWK 300  
DB |||||||



XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;  
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;  
XX WPI; 2003-468649/44.  
DR N-PSDB; ADN38917.  
XX  
XX Determining the presence or absence of a pathological cell in a patient,  
PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
PT a nucleic acid in a biological sample.  
XX  
XX Claim 12; SEQ ID NO 236; 1385pp; English.  
XX  
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
CC whose expression is upregulated or downregulated in specific cancers or  
CC other diseases such as angiogenic or fibrotic disorders, and to methods  
CC of determining the presence or absence of a pathological cell in a  
CC patient by detecting a nucleic acid at least 80% identical to those of  
CC the invention or by detecting a polypeptide of the invention. The  
CC invention also relates to expression vectors and host cells comprising a  
CC nucleic acid of the invention; antibodies which specifically bind a  
CC polypeptide of the invention; use of such antibodies for drug targeting;  
CC and methods of screening for modulators of activity or expression of the  
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,  
CC antibodies and methods are useful for diagnosing, prognosing and treating  
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
CC neovascularisation syndromes, scarring and uterine fibroids. They may  
CC also be useful in wound healing and in contraception. The present  
CC sequence represents a polypeptide of the invention.  
XX  
XX Sequence 472 AA;

Query Match 99.5%; Score 2412; DB 7; Length 472;  
Best Local Similarity 99.6%; Pred. No. 3.8e-240;  
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MXSILDLGADTTTPTITDLYVGSNDIQYEDIKGDMASKLGYFPQKPELTSFRGSPFOE 60  
DB 1 MXSILDLGADTTTPTITDLYVGSNDIQYEDIKGDMASKLGYFPQKPELTSFRGSPFOE 60  
QY 61 KMTAGNPPQVPADQVNIETFYNKSLSFFKNEENIQCGENFMDIECFWVLPNSQQLATA 120  
DB 61 KMTAGNPPQVPADQVNIETFYNKSLSFFKNEENIQCGENFMDIECFWVLPNSQQLATA 120  
QY 121 VLSLTGLTFTVLENLVLVILHSRSLRCRPSYHFTIGSLAVADLLGSLVIFVYSFIDFHFV 180  
DB 121 VLSLTGLTFTVLENLVLVILHSRSLRCRPSYHFTIGSLAVADLLGSLVIFVYSFIDFHFV 180  
QY 181 HRKDSRNVFLKGGVTSFTASVGLFLTAIARYISIHRLPLAYKRIIVTRPKAVAFCLM 240  
DB 181 HRKDSRNVFLKGGVTSFTASVGLFLTAIARYISIHRLPLAYKRIIVTRPKAVAFCLM 240  
QY 241 WTAIVAVLPLGWNCKLQSCVDSIFPHIDETFLMFWIGTVSVLLPLFVAYMYILNK 300  
DB 241 WTAIVAVLPLGWNCKLQSCVDSIFPHIDETFLMFWIGTVSVLLPLFVAYMYILNK 300  
QY 301 AHSHAVRMIQRGQTKSIIHTSDGKVQVTRPDQARMALRLAKTLVLVLVLIICWGPLL 360  
DB 301 AHSHAVRMIQRGQTKSIIHTSDGKVQVTRPDQARMALRLAKTLVLVLVLIICWGPLL 360  
QY 361 AIMYDVFGQWNLKIKTVFAFCMCLNLTVPPIIYALRSKDLRHFASMPSPSCGTAQ 420  
DB 361 AIMYDVFGQWNLKIKTVFAFCMCLNLTVPPIIYALRSKDLRHFASMPSPSCGTAQ 420  
QY 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEL 472  
DB 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEL 472

ABO60051  
ID ABO60051 standard; protein; 472 AA.  
XX  
AC ABO60051;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human genome derived single exon protein #6285.  
XX  
KW Human; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
PN US2003194704-A1.  
XX  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PR 03-APR-2002; 2002US-00029386.  
XX  
PI (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
WI; 2004-119264/12.  
XX  
PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
PS Claim 45; SEQ ID NO 33685; 80pp; English.  
XX  
CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 688 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subsequence, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe protein of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX



```
SQ Sequence 472 AA;
Query Match          99.5%; Score 2412; DB 8; Length 472;
Best Local Similarity 99.6%; Pred. No. 3.8e-240;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKSLDGLADTTFTTTTDLVYGSNDIQVEDIKGDMASKLGYPPQKPLTSFRGSPFQE 60
Db 1 MKSLDGLADTTFTTTTDLVYGSNDIQVEDIKGDMASKLGYPPQKPLTSFRGSPFQE 60

Qy 61 KMTAGDNPQLVPADQVNITEFYKSLSPFKENENIQCENFMDEICFVNLNPSQOLAIA 120
Db 61 KMTAGDNPQLVPADQVNITEFYKSLSPFKENENIQCENFMDEICFVNLNPSQOLAIA 120

Qy 121 VLSLTGLTFTVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSVIFYYSFIDFHFV 180
Db 121 VLSLTGLTFTVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSVIFYYSFIDFHFV 180

Qy 181 HRKDSRNVPFLKLGVTASTASVGSFLTAIARYISIHRLAYKRIVTRPKAVVAFCLM 240
Db 181 HRKDSRNVPFLKLGVTASTASVGSFLTAIARYISIHRLAYKRIVTRPKAVVAFCLM 240

Qy 241 WTIAIVAVPLLLGNCKEQSCDIFPHIDETFLMFWIGVTSVLLLFIVAYMYILWK 300
Db 241 WTIAIVAVPLLLGNCKEQSCDIFPHIDETFLMFWIGVTSVLLLFIVAYMYILWK 300

Qy 301 AHSNAVMIQRTGKSIITSEBGKQVQVTRPDQARMIRLAKTLVLLVLLICWGPILL 360
Db 301 AHSNAVMIQRTGKSIITSEBGKQVQVTRPDQARMIRLAKTLVLLVLLICWGPILL 360

Qy 361 AIMVYDVFGRKMKLIKTVFPCSMCLLNSVNPFIYALRSKDLRHAFRSPFCEGTAQ 420
Db 361 AIMVYDVFGRKMKLIKTVFPCSMCLLNSVNPFIYALRSKDLRHAFRSPFCEGTAQ 420

Qy 421 PLDNSMGDSDDLKHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472
Db 421 PLDNSMGDSDDLKHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 9
ADO29261
XX ID ADO29261 standard; protein; 472 AA.
XX AC ADO29261;
XX DT 29-JUL-2004 (first entry)
XX DE Human GPCR CNR1, SEQ ID NO:362.
XX KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW cytotatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianemic; antieborrheic;
KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW receptor.
XX OS Homo sapiens.
XX PN WO2004040000-A2.
XX PD 13-MAY-2004.
XX PF 09-SEP-2003; 2003WO-US028226.
XX
```

```
PR 09-SEP-2002; 2002US-0409303P.
XX 09-APR-2003; 2003US-0461329P.
PA (PRIM-) PRIMAL INC.
XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
XX WPI; 2004-390329/36.
DR N-PSDB; ADO29849.
XX
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
PS Claim 151; SEQ ID NO 362; 542pp; English.
XX
XX The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
CC invention. Note: The full sequence data for this patent did not form part
CC of the printed specification; those sequences not shown were obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 472 AA;
Query Match          99.5%; Score 2412; DB 8; Length 472;
Best Local Similarity 99.6%; Pred. No. 3.8e-240;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKSLDGLADTTFTTTTDLVYGSNDIQVEDIKGDMASKLGYPPQKPLTSFRGSPFQE 60
Db 1 MKSLDGLADTTFTTTTDLVYGSNDIQVEDIKGDMASKLGYPPQKPLTSFRGSPFQE 60

Qy 61 KMTAGDNPQLVPADQVNITEFYKSLSPFKENENIQCENFMDEICFVNLNPSQOLAIA 120
Db 61 KMTAGDNPQLVPADQVNITEFYKSLSPFKENENIQCENFMDEICFVNLNPSQOLAIA 120

Qy 121 VLSLTGLTFTVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSVIFYYSFIDFHFV 180
Db 121 VLSLTGLTFTVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSVIFYYSFIDFHFV 180

Qy 181 HRKDSRNVPFLKLGVTASTASVGSFLTAIARYISIHRLAYKRIVTRPKAVVAFCLM 240
Db 181 HRKDSRNVPFLKLGVTASTASVGSFLTAIARYISIHRLAYKRIVTRPKAVVAFCLM 240

Qy 241 WTIAIVAVPLLLGNCKEQSCDIFPHIDETFLMFWIGVTSVLLLFIVAYMYILWK 300
Db 241 WTIAIVAVPLLLGNCKEQSCDIFPHIDETFLMFWIGVTSVLLLFIVAYMYILWK 300
```

Db 241 WTIAIVIAVPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMILWK 300  
 Qy 301 AHSHAVRMIOQTQKSIITHTSDGKQVQVTRPDQARMAIRLAKTIVLILVLLIICWGPILL 360  
 Db 301 AHSHAVRMIOQTQKSIITHTSDGKQVQVTRPDQARMDIRLAKTIVLILVLLIICWGPILL 360  
 Qy 361 AIMVYDVFGKMKLIKTVPFACSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 Db 361 AIMVYDVFGKMKLIKTVPFACSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 Qy 421 PLDMSGDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 Db 421 PLDMSGDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

## RESULT 10

AEA81221  
 ID AEA81221 standard; protein; 472 AA.

AC AEA81221;

XX 25-AUG-2005 (first entry)

DT Human cannabinoid receptor 1 (brain) variant 1 protein.

DE screening; obesity; nutritional disorder; anorectic;

KW cannabinoid receptor 1.

KW

XX Homo sapiens.

OS US2005136465-A1.

PN 23-JUN-2005.

PD

XX 22-DEC-2004; 2004US-00019829.

XX 22-DEC-2003; 2003EP-00104902.

XX (CLER/) CLERC R G.

PA (DUCH/) DUCHATEAU-NGUYEN G.

PA (GARD/) GARDES C.

PA (MIZR/) MIZRAHI J.

PA (OSTE/) OSTENSON C.

XX

PI Clerc RG, Duchateau-Nguyen G, Gardes C, Mizrahi J, Ostenson C;

XX WPI; 2005-457507/46.

XX N-PSDB; AEA81160.

XX

XX Screening test compounds that reduce and/or prevent obesity involves

XX contacting cell expressing gene from alpha-two-glycoprotein.

XX Claim 18; SEQ ID NO 88; 21pp; English.

XX

XX The invention relates to a novel method for screening for test compounds

XX that reduce and/or prevent obesity. The method comprises contacting a

XX cell expressing a gene selected from SEQ ID NO. 1-12 or 25-85, with a

XX compound. The method of the invention demonstrates anorectic applications

XX and may be useful for screening for compounds that reduce and/or prevent

XX obesity. The current sequence is that of the human cannabinoid receptor 1

XX (brain) variant 1 protein of the invention. The sequence listing for the

XX specification can be located via the USPTO web-site.

XX

XX Sequence 472 AA;

XX

XX Query Match 99.5%; Score 2412; DB 9; Length 472;

XX Best Local Similarity 99.8%; Pred. No. 3.8e-240;

XX Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX

Qy 1 MKSILDLADTTTTRITDLYGSDNDIQVEDIKGDMASKLGYFPQKFLTSFRGSPFQ 60

Db 1 MKSILDLADTTTTRITDLYGSDNDIQVEDIKGDMASKLGYFPQKFLTSFRGSPFQ 60

Qy 61 KMTAGDNPOLVPADQVNITEFYNKSLSSFKENEENIQCGENFMDIECFMVLNPSQOLAIA 120  
 Db 61 KMTAGDNPOLVPADQVNITEFYNKSLSSFKENEENIQCGENFMDIECFMVLNPSQOLAIA 120  
 Qy 121 VLSLTGLTFTVLENLVLCVILHRSRLRCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHV 180  
 Db 121 VLSLTGLTFTVLENLVLCVILHRSRLRCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHV 180  
 Qy 181 HRKDSRNVPFLKGGVTASFTASVGLSLFLTATARYISIHRLPLAYKRIIVTRPKAVAFCLM 240  
 Db 181 HRKDSRNVPFLKGGVTASFTASVGLSLFLTADRIYISIHRLPLAYKRIIVTRPKAVAFCLM 240  
 Qy 241 WTIAIVIAVPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMILWK 300  
 Db 241 WTIAIVIAVPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMILWK 300  
 Qy 301 AHSHAVRMIOQTQKSIITHTSDGKQVQVTRPDQARMAIRLAKTIVLILVLLIICWGPILL 360  
 Db 301 AHSHAVRMIOQTQKSIITHTSDGKQVQVTRPDQARMDIRLAKTIVLILVLLIICWGPILL 360  
 Qy 361 AIMVYDVFGKMKLIKTVPFACSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 Db 361 AIMVYDVFGKMKLIKTVPFACSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 Qy 421 PLDMSGDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 Db 421 PLDMSGDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

## RESULT 11

ABB56338

ID ABB56338 standard; protein; 472 AA.

XX ABB56338;

XX 18-FEB-2002 (first entry)

XX Non-endogenous human GPCR protein, SEQ ID NO: 469.

XX Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;

XX constitutively activated GPCR; agonist; disease.

XX Homo sapiens.

XX Synthetic.

XX WO200177172-A2.

XX 18-OCT-2001.

XX 05-APR-2001; 2001WO-US011098.

XX 07-APR-2000; 2000US-0195747P.

XX (AREN-) ARENA PHARM INC.

XX Lehmann-Bruinsma K, Liauw CW, Lin I;

XX WPI; 2001-648759/74.

XX N-PSDB; ABI97974.

XX Identifying agonists of G protein-coupled receptors (GPCRs) for use in  
 PT disease treatment, comprises contacting candidate compounds with versions  
 PT of GPCRs.  
 XX Claim 1; Page 270-271; 394pp; English.  
 XX The invention relates to G protein-coupled receptors (GPCRs) for which  
 CC the endogenous ligand has been identified. Non-endogenous constitutively  
 CC activated versions of known GPCRs are used in the invention for the  
 CC direct identification of candidate compounds as receptor agonists,  
 CC inverse agonists or partial agonists. Such agonists are useful as  
 CC therapeutic agents for diseases or disorders associated with GPCRs. The





Db 301 AHSHAVRMIOQTGKSIHTSEDGKVQVTRPDQARMDIRLAKTLVLVLIIICWGPLL 360  
QY 361 AIWYDVFQGNKLIKTVFAFCMCLLNSTVNPPIIYALRSKDLRHAFRSMFPSCGTAQ 420  
Db 361 AIWYDVFQGNKLIKTVFAFCMCLLNSTVNPPIIYALRSKDLRHAFRSMFPSCGTAQ 420  
QY 421 PLDNSMGSDCLHKHANNAAVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
Db 421 PLDNSMGSDCLHKHANNAAVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 13  
ID AAR14195  
XX AAR14195 standard; protein; 473 AA.  
XX AC AAR14195;

XX 25-MAR-2003 (revised)  
DT 17-DEC-2001 (revised)  
DT 19-DEC-1991 (first entry)  
XX Rat cannabinoid receptor SKR6.  
XX Cannabis sativa; marijuana; drug test; substance K receptor.  
XX Rattus.

PH Key Location/Qualifiers  
FT Modified-site 78  
FT /label= OTHER  
FT /note= "N-glycosylated"  
FT Modified-site 84  
FT /label= OTHER  
FT /note= "N-glycosylated"  
FT Modified-site 113  
FT /label= OTHER  
FT /note= "N-glycosylated"  
FT Domain 118..143  
FT /label= hydrophobic\_domain\_I  
FT Domain 156..173  
FT /label= hydrophobic\_domain\_II  
FT Domain 189..213  
FT /label= hydrophobic\_domain\_III  
FT Domain 234..256  
FT /label= hydrophobic\_domain\_IV  
FT Domain 276..300  
FT /label= hydrophobic\_domain\_V  
FT Domain 346..366  
FT /label= hydrophobic\_domain\_VI  
FT Domain 379..400  
FT /label= hydrophobic\_domain\_VII

XX USN7564075-N.  
XX 03-SEP-1991.  
XX 08-AUG-1990; 90US-00564075.  
XX 08-AUG-1990; 90US-00564075.  
XX (USSH ) NAT INST OF HEALTH.  
XX (USDC ) US SEC OF COMMERCE.  
XX Matsuda L, Brownstein M;  
XX WPI; 1991-303326/41.  
XX DNA encoding mammalian cannabinoid receptor - used for producing receptor  
XX for screening drugs and ligands and in detection.  
XX Disclosure; Fig 1; 25pp; English.

XX SKR6 cDNA encoding the rat cannabinoid receptor was isolated from a rat

CC cerebral cortex cDNA library. An EcoRI-XbaI fragment was used to screen a  
CC human cosmid library and a positive clone was identified. There is ca. 97  
CC per cent homology between the deduced amino acid sequences of the human  
CC and rat cannabinoid receptors. The specification includes the sequence of  
CC a cDNA fragment which encodes this rat cannabinoid receptor protein but  
CC the copy quality is extremely poor. Recombinantly produced receptor can  
CC be used to screen for new drugs suitable for treatment of cannabinoid-  
CC treatable conditions, e.g. glaucoma, bronchial asthma, etc. See AAR14003  
CC and AAR14196 for human cannabinoid receptor sequences. (Note: Revised  
CC entry submitted to correct the patent number format of US Government-  
CC owned NTIS applications to prevent clashes with ongoing US granted patent  
CC numbers. For further information please visit the Derwent web site at  
CC www.derwent.com/dwpi/updates/ncis\_us.html.) (Updated on 25-MAR-2003 to  
CC correct PA field.)  
XX SQ Sequence 473 AA;

Query Match 96.7%; Score 2343.5; DB 2; Length 473;  
Best Local Similarity 96.8%; Pred. No. 4.7e-233;  
Matches 458; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
QY 1 MXSILDLGLADTTFRITITDLYVGSNDIQYEDIKGDMASKLGYFPQKFLTSFRGSPFQE 60  
Db 1 MXSILDLGLADTTFRITITDLYVGSNDIQYEDIKGDMASKLGYFPQKFLTSFRGSPFQE 60  
QY 61 KMTAGDNPOLVPA-DQVNITEFYFNKLSFSFKENEENIQCGENFMDECFMILNPSQOLAI 119  
Db 61 KMTAGDNPOLVPA-DQVNITEFYFNKLSFSFKENEENIQCGENFMDECFMILNPSQOLAI 120  
QY 120 AVLSLTLGFTVLENLVLCVILHSRLCRPSYHFIGSLAVADLLGSVIFYSFIDFHV 179  
Db 121 AVLSLTLGFTVLENLVLCVILHSRLCRPSYHFIGSLAVADLLGSVIFYSFIDFHV 180  
QY 180 FHRKDSNVFLFKLGVTASFTASVGSFLTAIARYISIHRLAYKRIVTRPKAVVAFCL 239  
Db 181 FHRKDSNVFLFKLGVTASFTASVGSFLTAIARYISIHRLAYKRIVTRPKAVVAFCL 240  
QY 240 MWTIAIVIAVPLLGWNCKLQSVCSDFPHIDETVLMFVIGTSLVLLFIVAYMYILW 299  
Db 241 MWTIAIVIAVPLLGWNCKLQSVCSDFPHIDETVLMFVIGTSLVLLFIVAYMYILW 300  
QY 300 KAHSHAVRMIOQTGKSIHTSEDGKVQVTRPDQARMAIRLAKTLVLVLIIICWGPL 359  
Db 301 KAHSHAVRMIOQTGKSIHTSEDGKVQVTRPDQARMAIRLAKTLVLVLIIICWGPL 360  
QY 360 LAIMYDVFQGNKLIKTVFAFCMCLLNSTVNPPIIYALRSKDLRHAFRSMFPSCGTA 419  
Db 361 LAIMYDVFQGNKLIKTVFAFCMCLLNSTVNPPIIYALRSKDLRHAFRSMFPSCGTA 420  
QY 420 QPLDNSMGSDCLHKHANNAAVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
Db 421 QPLDNSMGSDCLHKHANNAAVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 473

RESULT 14  
ADD46381  
ID ADD46381 standard; protein; 473 AA.  
XX AC ADD46381;  
XX 29-JAN-2004 (first entry)  
XX Rat Protein P20272, SEQ ID NO 12061.  
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
XX Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX Rattus norvegicus.  
XX WO2003016475-A2.  
XX 27-FEB-2003.



Db 1 MKSILDGLADTTFRITTTDLLVYGSNDIQYEDIKGDMSKLGYPQKFPFLTFRGSPFOE 60  
Qy 61 KMTAGNPOLVPA-DQVNIITEFYNKSLSSFKENEENIQCGENFMDIECFWNLNPSOQLAI 119  
Db 61 KMTAGNSPLVPAGDTTNIITEFYNKSLSSFKENEDNIQCGENFMDIECFWNLNPSOQLAI 120  
Qy 120 AVLSLTGTFVLENLVLCVILHRSRLRCRPSYHFIGSLAVADLLGSVIFVYSFIDPHV 179  
Db 121 AVLSLTGTFVLENLVLCVILHRSRLRCRPSYHFIGSLAVADLLGSVIFVYSFVDFHV 180  
Qy 180 FHRKDSRNVLFLKLGVTASFTASVGSFLTAIARYISIHRLAYKRIVTRPKAVVAFCL 239  
Db 181 FHRKDSPNVFLKLGVTASFTASVGSFLTAIDRYISIHRLAYKRIVTRPKAVVAFCL 240  
Qy 240 MMTIAIVIAVLPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVYAYMYILW 299  
Db 241 MMTIAIVIAVLPLLGWNCCKLQSVCSDFPLIDETVLMFWIGVTSVLLLFIVYAYMYILW 300  
Qy 300 KAHSHAVRMIQRGTQKSIITHTSEDKGVQVTRPDQARMAIRLAKTLVLILVLLIICWGEL 359  
Db 301 KAHSHAVRMIQRGTQKSIITHTSEDKGVQVTRPDQARMDIRLAKTLVLILVLLIICWGEL 360  
Qy 360 LAIMVYDVFQKXNKLIKTVPFCSMLCLLNSTVNPITIALRSKDLRHAFRSMFPSCGTA 419  
Db 361 LAIMVYDVFQKXNKLIKTVPFCSMLCLLNSTVNPITIALRSKDLRHAFRSMFPSCGTA 420  
Qy 420 QPLDMSMGSDCLHKHANNAAHVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
Db 421 QPLDMSMGSDCLHKHANNAAHVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 473

Search completed: January 6, 2006, 21:08:29  
Job time : 194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 21:08:38 ; Search time 13 Seconds

(without alignments)  
308.138 Million cell updates/sec

Title: US-10-521-420-1

Perfect score: 2424

Sequence: 1 MSLDGLADTTTTRITTDL.....TVKIAKVTMSVSDTSAEAL 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 61072 seqs, 8486849 residues

Total number of hits satisfying chosen parameters: 61072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /cgn2\_5/ptodata/2/pubpaa/US08\_NEW\_PUB\_PEP.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PEP.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB\_PEP.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2412	99.5	472	7	US-11-127-877-39
2	710	29.3	360	6	US-10-851-667A-26
3	363	15.0	364	7	US-11-067-884-2
4	352.5	14.5	378	6	US-10-995-561-891
5	351.5	14.5	382	7	US-11-124-368A-173
6	351.5	14.5	382	7	US-11-124-368A-174
7	351.5	14.5	382	7	US-11-127-877-58
8	341	14.1	353	7	US-11-067-884-6
9	311.5	12.9	351	7	US-11-067-884-4
10	310.5	12.8	334	7	US-11-110-274-6
11	289	11.9	330	7	US-11-110-274-4
12	283	11.7	429	7	US-11-127-877-51
13	283	11.7	466	7	US-11-127-877-50
14	280	11.6	362	7	US-11-110-274-5
15	272	11.2	360	6	US-10-995-561-547
16	248.5	10.3	254	6	US-10-055-877-248
17	248.5	10.3	254	6	US-10-055-877-327
18	248.5	10.3	254	6	US-10-055-877-340
19	235.5	9.7	797	6	US-10-995-561-802
20	231.5	9.6	317	6	US-10-995-561-798
21	228	9.4	318	7	US-11-127-877-38
22	222.5	9.2	480	6	US-10-521-162-40
23	221	9.1	259	6	US-10-055-877-225
24	221	9.1	259	6	US-10-055-877-237
25	216.5	8.9	471	6	US-10-995-561-901

26	213	8.8	365	6	US-10-875-716-9	Sequence 9, Appli
27	209	8.6	384	7	US-11-080-991-26	Sequence 26, Appl
28	201	8.3	412	7	US-11-170-153-8	Sequence 8, Appli
29	201	8.3	412	7	US-11-170-166-8	Sequence 8, Appli
30	201	8.3	412	7	US-11-170-351-8	Sequence 8, Appli
31	201	8.3	412	7	US-11-170-153-6	Sequence 6, Appli
32	201	8.3	415	7	US-11-170-166-6	Sequence 6, Appli
33	201	8.3	415	7	US-11-170-351-6	Sequence 6, Appli
34	198	8.2	407	7	US-11-127-877-56	Sequence 56, Appl
35	197	8.1	412	7	US-11-170-153-4	Sequence 4, Appli
36	197	8.1	412	7	US-11-170-166-4	Sequence 4, Appli
37	197	8.1	412	7	US-11-170-351-4	Sequence 2, Appli
38	197	8.1	415	7	US-11-170-153-2	Sequence 2, Appli
39	197	8.1	415	7	US-11-170-166-2	Sequence 2, Appli
40	197	8.1	415	7	US-11-170-351-2	Sequence 2, Appli
41	195	8.0	400	7	US-11-127-877-55	Sequence 55, Appl
42	195	8.0	415	7	US-11-127-877-45	Sequence 45, Appl
43	194.5	8.0	757	6	US-10-055-877-157	Sequence 157, App
44	194	8.0	712	6	US-10-521-162-12	Sequence 12, Appl
45	193	8.0	1115	6	US-10-055-877-160	Sequence 160, App

ALIGNMENTS

RESULT 1

US-11-127-877-39  
; Sequence 39, Application US/11127877  
; Publication No. US20050287565A1  
; GENERAL INFORMATION:  
; APPLICANT: Merchiers, Pascal G.  
; APPLICANT: Hoffmann, Marcel  
; APPLICANT: Spittaels, Koenraad F. F.  
; APPLICANT: Laenen, Wendy  
; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting  
; FILE REFERENCE: P27,800-B USA  
; CURRENT APPLICATION NUMBER: US/11/127,877  
; CURRENT FILING DATE: 2005-05-12  
; PRIOR APPLICATION NUMBER: 60/570,352  
; PRIOR FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: 60/603,948  
; PRIOR FILING DATE: 2004-08-24  
; NUMBER OF SEQ ID NOS: 590  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 39  
; LENGTH: 472  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-127-877-39

Query Match	99.5%	Score	2412	DB	7	Length	472
Best Local Similarity	99.6%	Pred. No.	6e-199				
Matches	470	Conservative	0	Mismatches	2	Indels	0
Gaps	0						
Qy	1	MKSILDLGLADTTTTRITTTDLYVGSNDIQVEDIKGDWASKLGYFPKPLTSFRGSPQ	60				
Db	1	MKSILDLGLADTTTTRITTTDLYVGSNDIQVEDIKGDWASKLGYFPKPLTSFRGSPQ	60				
Qy	61	KMTAGDNPQLVPADQVNITEFYKSLSSFKENENIQCGENFMDECFMVLNPSQOLAIA	120				
Db	61	KMTAGDNPQLVPADQVNITEFYKSLSSFKENENIQCGENFMDECFMVLNPSQOLAIA	120				
Qy	121	VLSLTGLTFTVLENLVLCVILHRSRUCRPSYHFIGSLAVADLLGVSIVFYSPDFHVF	180				
Db	121	VLSLTGLTFTVLENLVLCVILHRSRUCRPSYHFIGSLAVADLLGVSIVFYSPDFHVF	180				
Qy	181	HRKDSRVFLKLGVTASFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM	240				
Db	181	HRKDSRVFLKLGVTASFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM	240				
Qy	241	WTIAIVAVLFLGNCKLQSCVDIPPHIDEITVLMFVIGVTSVLLLFIVTAYNYILWK	300				
Db	241	WTIAIVAVLFLGNCKLQSCVDIPPHIDEITVLMFVIGVTSVLLLFIVTAYNYILWK	300				

```
Db 241 WTIAIVAVLPGLGWNCEKLSQVCSDFPHIDETVLMFWIGVTSVLLLFIVYAYMYLWK 300
Qy 301 AHSHAVRMIOQTQKSIHHTSDGKVQVTRPDQARMAJRLAKTLVLILVLIICWGPLL 360
Db 301 AHSHAVRMIOQTQKSIHHTSDGKVQVTRPDQARMDIRLAKTLVLILVLIICWGPLL 360
Qy 361 AIMVYDVGGMNKLKTVFAFCMCLLNSTVNPPIIYALRSKDLRHAFRSMFPSCBGTQ 420
Db 361 AIMVYDVGGMNKLKTVFAFCMCLLNSTVNPPIIYALRSKDLRHAFRSMFPSCBGTQ 420
Qy 421 PLDNSMGDSCLHKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472
Db 421 PLDNSMGDSCLHKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 2
US-10-851-667A-26
; Sequence 26, Application US/10851667A
; Publication No. US20050260608A1
; GENERAL INFORMATION:
; APPLICANT: Zimmer, Andreas
; APPLICANT: Karsak, Meliha
; APPLICANT: de Vernejoul, Marie-Christine
; APPLICANT: Bab, Itai
; APPLICANT: Shohami, Esther
; APPLICANT: Mechoulam, Raphael
; TITLE OF INVENTION: DELAYING ONSET OF, PREVENTING AND/OR TREATING OSTEOPOROSIS
; FILE REFERENCE: 28030
; CURRENT APPLICATION NUMBER: US/10/851,667A
; CURRENT FILING DATE: 2004-05-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (63)..(63)
; OTHER INFORMATION: Non-synonymous polymorphism: Gln or Arg
US-10-851-667A-26

Query Match 29.3%; Score 710; DB 6; Length 360;
Best Local Similarity 45.2%; Pred. No. 1.2e-53;
Matches 149; Conservative 57; Mismatches 94; Indels 30; Gaps 6;

Qy 78 ITFYNKSLSSFKENEENIQCGENFMDIECFWVLPNSQQLAIAVLSLTGTFTVLENLIV 137
Db 6 VTEIANGSKDGLDSN-----PMKDYMLISGPQKTAVALCTLLGLLSALENAV 54
Qy 138 LCVILHSRLRCPSPSYHFTGSLAVADLLGVSIVFYSDIFPHVPHRDXSRNVFLFKLGVT 197
Db 55 LYLILSSHLLRKPSPYLFITGSLAGADFLASWFACFSFVPHVPHGVDKSAVFLKIGSVT 114
Qy 198 ASPTASVGSFLTAIARYSIHHRPLAYKRVTRPKAVAFCLMTIAIVAVLPGLGWN 257
Db 115 MPTASVGSLLTAIDRYLCRLRPPSYKALLTFGRALVTGIMVWLSALVSILPLMGWTC 174
Qy 258 EKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVYAYMYLWKASHAVRMIOQTQKSI 317
Db 175 --CPRCESEFLPIPNDYLLSWLLFTAFLPSTGIITYYGHVWKAHQHVA-----SL 223
Qy 318 IHTSDGKVQVTRPDQARMA--IRLAKTLVLILVLIICWGPLLAIMVYDVGGMNKL 375
Db 224 SGH--QDRQV-----PGVARRMLDRLAKTLGLVLAVLLICWFPVLMAHSLATLSDQV 277
Qy 376 KTVFAFCMCLLNSTVNPPIIYALRSKDLR 405
Db 278 KKAFAFCMCLLNSTVNPPIIYALRSGEIR 307

RESULT 3
```

```
US-11-067-884-2
; Sequence 2, Application US/11067884
; Publication No. US20050261252A1
; GENERAL INFORMATION:
; APPLICANT: Miller, Duane D.
; APPLICANT: Tigyi, Gabor
; APPLICANT: Dalton, James T.
; APPLICANT: Sardar, Vineet M.
; APPLICANT: Elrod, Don B.
; APPLICANT: Xu, Huiping
; APPLICANT: Baker, Daniel L.
; APPLICANT: Wang, Dean
; APPLICANT: Lillom, Karoly
; APPLICANT: Fischer, David J.
; APPLICANT: Virag, Tamas
; APPLICANT: Nusser, Nora
; TITLE OF INVENTION: LPA RECEPTOR AGONISTS AND ANTAGONISTS AND METHODS OF
; TITLE OF INVENTION: USE
; FILE REFERENCE: 20609/305
; CURRENT APPLICATION NUMBER: US/11/067,884
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 60/190,370
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 09/811,838
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-067-884-2

Query Match 15.0%; Score 363; DB 7; Length 364;
Best Local Similarity 27.2%; Pred. No. 5.5e-24;
Matches 108; Conservative 72; Mismatches 167; Indels 50; Gaps 10;

Qy 59 QEKMTAGNPQLVPADQVNIETFYNKSLSSFKENEENIQCGENFMDIECFWVLPNSQQLA 118
Db 13 QPQFTAMNEPQC-----FYNESIAFFYNNRSGKHLATE-----WNTVSKLV 52
Qy 119 IAVLSLTGTFTVLENLILVLCVILHSRLRCPSPHYFTGSLAVADLLGVSIVFYSDIFDH 178
Db 53 MG-LGITVCIFTMLANLLVMVAIYVNRFRHP-FIYLMANLAAADFFAGLAYFYLMNTG 110
Qy 179 VFHRKDSRVFLFKLGGVTFASVGSFLTAIARYSIHHRPLAYKRVITRPKAVVAF 238
Db 111 PNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAERHITVFRMQLHTRMSNR-RVVVVIV 169
Qy 239 LMWTIAIVAVLPGLGWNCEKLSQVCSDFPHIDETVLMFWIGVTSVLLLFIVYAYMYL 298
Db 170 VIWMTAIVMGAIPSGWNCICDIENCSNMAPLYSDSYLVFAIFNLVTFVWVWVLYAHIF 229
Qy 299 WKAHSHAVRMIOQTQKSIHHTSDGKVQVTRPDQARMAIRLAKTLVLILVLIICWGP 358
Db 230 GYVORTWEMSR-----HSSGPRNRDT-----MMSLLKTVVIVLGAFLIICWTP 273
Qy 359 LLAIWYDVFGGMNKLKTVFAFCMCLLNSTVNPPIIYALRSKDLRHAFRSMFPSCBGT 418
Db 274 GLVLLLDVCCQCDVL-AYEKFFLLLAEFNANMNPPIIYSYRDKEMSAIFRQIL-----C 327
Qy 419 AQLDNSMGDSCLHKHANNA--AASVHRAAESCI 450
Db 328 QRSENPQTGPTSSDRSASSLNHTILAGVHSDHVS 364

RESULT 4
US-10-995-561-891
; Sequence 891, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
```

```
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 891
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-891

Query Match      14.5%; Score 352.5; DB 6; Length 378;
Best Local Similarity 29.7%; Pred. No. 4.6e-23;
Matches 105; Conservative 61; Mismatches 131; Indels 57; Gaps 13;

QY 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFIFGSLAVADLLGVSFVIFDHFV 180
Db 44 VLFVICSFVLENLVLCVILHSRLCRPSYHFIFGSLAVADLLGVSFVIFDHFV 99
QY 181 HRKD---SRNVFLFKGGVTSFTASVGSFLFTRAIARYIS--HRPL-AKRVITRPKAV 234
Db 100 GKTFSLSPVWFLREGSMFVALGASTCSLLAIAERHLTWIKMRPYDANKR--HRVFL 157
QY 235 VAFCLMTIAIVTLVLLGNCKEQLQSCSDIFPHIDETVLMFWIGVTSVLLFIV--Y 292
Db 158 IGMG--WLIATFLGALPILGNLHNLPCDSTILPLYSKKYIAFCISFTAILVTIVILY 215
QY 293 AYVILWKAHSHAVRMIRGTQKSIITHTSDGKQVQVTRPDQARMAIRLAKTLVLLV 352
Db 216 ARIYFLVKSRR-----KVNHNRSERSMALLRTVIVVSFV 252
QY 353 IICWGPLLAIMVYDVGKQNK---LITVFAPCSMLCLLNSTVNPITIALRSKDLRAFR 409
Db 253 IACWSPLFILIDVACRVQACPIFLKAQWFI--VLAVLNSAMNPVITVLASKEMRRAFF 310
QY 410 SMFPSC-----EGTAQPLDMSMGSDCLHKHANNAS-----VHRAAESCI 450
Db 311 RLVCNCLVRGRGARSPIQALDPSRKSSSSSNNSSHSPKVKEDLPHTAPSSCI 364

RESULT 5
US-11-124-368A-173
; Sequence 173, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-173

Query Match      14.5%; Score 351.5; DB 7; Length 382;
Best Local Similarity 27.6%; Pred. No. 5.7e-23;
Matches 105; Conservative 67; Mismatches 125; Indels 83; Gaps 13;

QY 68 POLVP-----ADQVN---ITEFYN-----KSLSSPKENEENIQCGENFMDIECFMVL 111
Db 3 PTSVPLVKAHRSVSDVYVYDIIVRHNYTGLKLNISADKENSIKL----- 47
QY 112 NPSQQLAIAVLSTLTGTFVLENLVLCVILHSRLCRPSYHFIFGSLAVADLLGVSFV 171
Db 48 -----TSVVFLICCFIILENIFVLLTIWTKKFF-RPMYFYIGNLALSLLAGVAYT 99
QY 172 YSFIDFVHRKDSRNVLFLKGGVTSFTASVGSFLFTRAIARYIS-----IHRPLAYKR 226
Db 100 ANLLSGATTYKLTTPAQWFLREGSMFVALSASFVSLAIAERYITMLKMKLHNSNNFR 159
QY 227 IVTRPKAVVAPCLMWTIAIVTLVLLGNCKEQLQSCSDIFPHIDETVLMFWIGVTSV 286
Db 160 LF-----LLISAC--WVLSILGGLPIMGWNCISALSSCSVTLPLVYHKHYILFCTVTFTLL 213
QY 287 LLFIV--YAYVILWKAHSHA-----VRMIQRGTQKSIITHTSDGKQVQVTRPDQARMA 338
Db 214 LLSIVILYCRFYSLVTRSRRLFRKNISKASRSSEKS----- 251
QY 339 IRLAKTLVLLVLLIICWGPLLAIMVYDVGKQNKLIKTVP--AFCSMLCLLNSTVNP 396
Db 252 LALLKTVIIVLSVFIACWAPLFTILLDVGKV-KTCDILFRABYFLVLAVLNSGTNP 310
QY 397 YALRSKDLRAFRSMFPSC 416
Db 311 YTLTNKEMRRAFIRIMSCCK 330

RESULT 6
US-11-124-368A-174
; Sequence 174, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-174

Query Match      14.5%; Score 351.5; DB 7; Length 382;
Best Local Similarity 27.6%; Pred. No. 5.7e-23;
Matches 105; Conservative 67; Mismatches 125; Indels 83; Gaps 13;

QY 68 POLVP-----ADQVN---ITEFYN-----KSLSSPKENEENIQCGENFMDIECFMVL 111
Db 3 PTSVPLVKAHRSVSDVYVYDIIVRHNYTGLKLNISADKENSIKL----- 47
QY 112 NPSQQLAIAVLSTLTGTFVLENLVLCVILHSRLCRPSYHFIFGSLAVADLLGVSFV 171
Db 48 -----TSVVFLICCFIILENIFVLLTIWTKKFF-RPMYFYIGNLALSLLAGVAYT 99
QY 172 YSFIDFVHRKDSRNVLFLKGGVTSFTASVGSFLFTRAIARYIS-----IHRPLAYKR 226
Db 100 ANLLSGATTYKLTTPAQWFLREGSMFVALSASFVSLAIAERYITMLKMKLHNSNNFR 159
QY 227 IVTRPKAVVAPCLMWTIAIVTLVLLGNCKEQLQSCSDIFPHIDETVLMFWIGVTSV 286
Db 160 LF-----LLISAC--WVLSILGGLPIMGWNCISALSSCSVTLPLVYHKHYILFCTVTFTLL 213
QY 287 LLFIV--YAYVILWKAHSHA-----VRMIQRGTQKSIITHTSDGKQVQVTRPDQARMA 338
```

```
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 891
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-891

Query Match      14.5%; Score 352.5; DB 6; Length 378;
Best Local Similarity 29.7%; Pred. No. 4.6e-23;
Matches 105; Conservative 61; Mismatches 131; Indels 57; Gaps 13;

QY 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFIFGSLAVADLLGVSFVIFDHFV 180
Db 44 VLFVICSFVLENLVLCVILHSRLCRPSYHFIFGSLAVADLLGVSFVIFDHFV 99
QY 181 HRKD---SRNVFLFKGGVTSFTASVGSFLFTRAIARYIS--HRPL-AKRVITRPKAV 234
Db 100 GKTFSLSPVWFLREGSMFVALGASTCSLLAIAERHLTWIKMRPYDANKR--HRVFL 157
QY 235 VAFCLMTIAIVTLVLLGNCKEQLQSCSDIFPHIDETVLMFWIGVTSVLLFIV--Y 292
Db 158 IGMG--WLIATFLGALPILGNLHNLPCDSTILPLYSKKYIAFCISFTAILVTIVILY 215
QY 293 AYVILWKAHSHAVRMIRGTQKSIITHTSDGKQVQVTRPDQARMAIRLAKTLVLLV 352
Db 216 ARIYFLVKSRR-----KVNHNRSERSMALLRTVIVVSFV 252
QY 353 IICWGPLLAIMVYDVGKQNK---LITVFAPCSMLCLLNSTVNPITIALRSKDLRAFR 409
Db 253 IACWSPLFILIDVACRVQACPIFLKAQWFI--VLAVLNSAMNPVITVLASKEMRRAFF 310
QY 410 SMFPSC-----EGTAQPLDMSMGSDCLHKHANNAS-----VHRAAESCI 450
Db 311 RLVCNCLVRGRGARSPIQALDPSRKSSSSSNNSSHSPKVKEDLPHTAPSSCI 364

RESULT 5
US-11-124-368A-173
; Sequence 173, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-173

Query Match      14.5%; Score 351.5; DB 7; Length 382;
Best Local Similarity 27.6%; Pred. No. 5.7e-23;
Matches 105; Conservative 67; Mismatches 125; Indels 83; Gaps 13;

QY 68 POLVP-----ADQVN---ITEFYN-----KSLSSPKENEENIQCGENFMDIECFMVL 111
Db 3 PTSVPLVKAHRSVSDVYVYDIIVRHNYTGLKLNISADKENSIKL----- 47
QY 112 NPSQQLAIAVLSTLTGTFVLENLVLCVILHSRLCRPSYHFIFGSLAVADLLGVSFV 171
Db 48 -----TSVVFLICCFIILENIFVLLTIWTKKFF-RPMYFYIGNLALSLLAGVAYT 99
QY 172 YSFIDFVHRKDSRNVLFLKGGVTSFTASVGSFLFTRAIARYIS-----IHRPLAYKR 226
Db 100 ANLLSGATTYKLTTPAQWFLREGSMFVALSASFVSLAIAERYITMLKMKLHNSNNFR 159
QY 227 IVTRPKAVVAPCLMWTIAIVTLVLLGNCKEQLQSCSDIFPHIDETVLMFWIGVTSV 286
Db 160 LF-----LLISAC--WVLSILGGLPIMGWNCISALSSCSVTLPLVYHKHYILFCTVTFTLL 213
QY 287 LLFIV--YAYVILWKAHSHA-----VRMIQRGTQKSIITHTSDGKQVQVTRPDQARMA 338
```



```
Db 214 LLSIVLYCRIYSLVTRSRRLTRFRNKISKASRSSEKS----- 251
QY 339 IRLAKTLVLLVLLIICWGPLLAIMVYDFGKNNKLIKTVF--AFCSMLCLLNSTVNPPII 396
Db 252 LALKTVIIVLSVFIACWAPLFLILLDDVGCKV-KTCDILFRAEYFLVLAVLNSGTNPPII 310
QY 397 YALRSKDLRHAFRSMPPSC 416
Db 311 YTLTNKEMRAFRINSCCK 330

RESULT 7
US-11-127-877-58
; Sequence 58, Application US/11127877
; Publication NO. US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merckhiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koenraad F. F.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27,800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877
; CURRENT FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 590
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 58
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-877-58

Query Match 14.5%; Score 351.5; DB 7; Length 382;
Best Local Similarity 27.6%; Pred. No. 5.7e-23;
Matches 105; Conservative 67; Mismatches 125; Indels 83; Gaps 13;

QY 68 PQLVP-----ADQVN---ITBFYN-----KSLSSFKEENENIOGGENFMDIECFMVL 111
Db 3 PTVPLVKAHRSSVDVNYDIIVRHVNYTGKLNISADKENSIKL----- 47
QY 112 NPSQQLAIAVLSLTGLTFTVLENLVLCVILHRSRLCRPSYHFIGSLAVADLLGSVIFV 171
Db 48 -----TSVFTLICCFILENFVLLTTWTKXPH-RPMYYFIGNLALSDLLAGVAYT 99
QY 172 YSFIDFVHRKDSRVNLFPLKGVTSFTASVGSFLFTAIAIYIS-----IHRPLAYKR 226
Db 100 ANLLSGATTYKLTTPAQWREGSMFVALSVFSLTAIAIERIITMLKMKHGNGNNFR 159
QY 227 IVTRPRKAVAFCLMWTIAIVLPLLGWNCCKLQSVCSDFPHIDETVLMFVIGTSVL 286
Db 160 LF-----LLISAC--WISLILGSLPGWNCISALSSCSVLPYKHXYILFCTTVFTLL 213
QY 287 LLFIV--YAWMYILWKAHSHA-----VRMIQRTQKSIHHTSBDGKVQVTRPDQARMA 338
Db 214 LLSIVLYCRIYSLVTRSRRLTRFRNKISKASRSSEKS----- 251
QY 339 IRLAKTLVLLVLLIICWGPLLAIMVYDFGKNNKLIKTVF--AFCSMLCLLNSTVNPPII 396
Db 252 LALKTVIIVLSVFIACWAPLFLILLDDVGCKV-KTCDILFRAEYFLVLAVLNSGTNPPII 310
QY 397 YALRSKDLRHAFRSMPPSC 416
Db 311 YTLTNKEMRAFRINSCCK 330

RESULT 8
US-11-067-884-6
```

```
; Sequence 6, Application US/11067884
; Publication NO. US20050261252A1
; GENERAL INFORMATION:
; APPLICANT: Miller, Duane D.
; APPLICANT: Tigyi, Gabor
; APPLICANT: Dalton, James T.
; APPLICANT: Sardar, Vineet M.
; APPLICANT: Elrod, Don B.
; APPLICANT: Xu, Huiping
; APPLICANT: Baker, Daniel L.
; APPLICANT: Wang, Dean
; APPLICANT: Lilloom, Karoly
; APPLICANT: Fischer, David J.
; APPLICANT: Nusseer, Nora
; TITLE OF INVENTION: LPA RECEPTOR AGONISTS AND ANTAGONISTS AND METHODS OF
; TITLE OF INVENTION: USE
; FILE REFERENCE: 20609/305
; CURRENT APPLICATION NUMBER: US/11/067,884
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 60/190,370
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 09/811,838
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-067-884-6

Query Match 14.1%; Score 341; DB 7; Length 353;
Best Local Similarity 26.9%; Pred. No. 4e-22;
Matches 80; Conservative 68; Mismatches 123; Indels 26; Gaps 6;

QY 121 VLSLTGLTFTVL-----ENLLVLCVILHRSRLCRPSYHFIGSLAVADLLGSVIFVYSFID 176
Db 31 VIVLCVGTFFCLFIFFNSLSVIAAVIKNRKHF-PFYLLANLAADFFAGIAYVFLMFN 89
QY 177 FHVFRKDSRVNLFPLKGVTSFTASVGSFLFTAIAIYIS-IHRPLAYKRIVTRPKAVVA 236
Db 90 TGPVSKTLLTNVNFURQGLDSSLTASLTSLNLLVIAVERHWSIMRMKVHNSL-TKKRVTL 148
QY 237 FCLMTTIAIVLPLLGWNCCKLQSVCSDFPHIDETVLMFVI--GVTSVLLLFIVYAY 294
Db 149 ILLVMAIAIFMGAVPTLGNCLCNISACSLAPIYSRSLVFWTVSNLMAFLIMVVYL 208
QY 295 MYLWKAHSHAVRMIOQTQKSIHHTSBDGKVQVTRPDQARMAIRLAKTLVLLVLLI 354
Db 209 IYVYVK-----RKTNVLSPHTSGSI-----SRRRTPMKLMKMTVMTVLGAFV 250
QY 355 CWGPLLAIMVYDFGKNNKLIKTVFAPCSMLCLLNSTVNPPIIYALRSKDLRHAFRSM 411
Db 251 CWTPLGVLLLDGLNCRQCGVQHVKRWFLLLLALLNSVVPNPIIYSYKDEDMYGTMKKM 307

RESULT 9
US-11-067-884-4
; Sequence 4, Application US/11067884
; Publication NO. US20050261252A1
; GENERAL INFORMATION:
; APPLICANT: Miller, Duane D.
; APPLICANT: Tigyi, Gabor
; APPLICANT: Dalton, James T.
; APPLICANT: Sardar, Vineet M.
; APPLICANT: Elrod, Don B.
; APPLICANT: Xu, Huiping
; APPLICANT: Baker, Daniel L.
; APPLICANT: Wang, Dean
; APPLICANT: Lilloom, Karoly
; APPLICANT: Fischer, David J.
; APPLICANT: Virag, Tamas
```





```
QY 411 MFPSCEGTQAPLDNSMGDS 430
Db 311 VCCCCSSKTPF-RSRSPD 329

RESULT 12
US-11-127-877-51
; Sequence 51, Application US/11127877
; Publication No. US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merckhiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koenraad F. F.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27,800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877
; PRIOR FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 590
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-877-51

Query Match 11.7%; Score 283; DB 7; Length 429;
Best Local Similarity 26.2%; Pred. No. 4.7e-17;
Matches 95; Conservative 72; Mismatches 156; Indels 40; Gaps 12;

QY 111 LNPSQOLAIAVLSLTGTFTVLENLVLVCVILHSRLCRPSYHFTGSLAVAD-LLGSVI 169
Db 21 VNISKAILLGVILGGLILFGVLGNILVILSVACHRHLLH-SVTHYYIVNLAVADLLTSTV 79

QY 170 FVYSFIDFHVHRKDSRNVLFLKGGV-TASFTASVGSILFLTAIARYISIHRLAYKRV 228
Db 80 LPFSAL-PEVLGYWAFGRFCNIAAADVLCCTASIMGLCIISIDRYIGSVYPLRYPTIV 138

QY 229 TRPKAVAFCLMWTIAIVAVLPLLGWN--CEKLSQVCS-----DIPPHIDETVLMFW 279
Db 139 TORRGLMALLCWALSLVISIGLFGWRQPAPEDEITCQINEEPGVLFPSALGFSYL--- 195

QY 280 IGTVSVLLFIVYAYMYILKASHAVRMQR-----GTQKSIHHTSE---DGKQVQVTR 331
Db 196 ----PLAILVMYCRYVYVAKRESGLKGLTKDSQVTLRIHRKNAPAGGSGMASA 251

QY 332 PQQARMAIRL-----AKTLVLILVVLIIICWGPLLAIMVDVFGKKNKLIKTVFAPC 382
Db 252 KTKTHFSVRLKFSREKKAATLGIIVGCVFLCWLFPFLVMPIGSFPPDFKPSSETVFKIV 311

QY 383 SMLCLLNSTVNPITIALRSKDLRHAFRSMFP-SCGTAQPLDNSMGDSCLHKHANAAS 441
Db 312 FWLGYLNSCINPIIYPCSSQEFKAFQNVLRIOCLRRKQSSKHALGYT--LHPPSOAVEG 369

QY 442 VHR 444
Db 370 QHK 372

RESULT 14
US-11-110-274-5
; Sequence 5, Application US/11110274
; Publication No. US20050266502A1
; GENERAL INFORMATION:
; APPLICANT: Merckhiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koenraad F. F.
; TITLE OF INVENTION: Methods, Compositions and Compound Assays for Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27,697-A USA
; CURRENT APPLICATION NUMBER: US/11/110,274
; CURRENT FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: US 60/563,661
; PRIOR FILING DATE: 2004-04-20
; NUMBER OF SEQ ID NOS: 620
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-110-274-5
```



**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 21:01:17 ; Search time 231 Seconds  
(without alignments)

1441.599 Million cell updates/sec

Title: US-10-521-420-1

Perfect score: 2424

Sequence: 1 MKSILGLADTTFTTTDL.....TVKIAKVTMSVSDTDSAEAL 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2412	99.5	472	1	CNR1_HUMAN
2	2412	99.5	472	1	CNR1_PANTR
3	2412	99.5	472	2	Q4PLI4_HUMAN
4	2412	99.5	472	2	Q71SP5_MACMU
5	2380	98.2	472	2	Q506J9_HUMAN
6	2377	98.1	471	2	Q4VBM6_HUMAN
7	2343	96.7	473	1	CNR1_RAT
8	2340	96.6	473	1	CNR1_MOUSE
9	2340	96.6	473	2	Q5SF33_MOUSE
10	2335	96.3	472	1	CNR1_FELCA
11	2250	92.8	473	1	CNR1_PORGU
12	2209	91.2	439	2	Q5UB37_HUMAN
13	2064	85.2	473	1	CNR1_TARGR
14	2031	83.8	470	2	Q80LM1_XENLA
15	1954	80.6	411	2	Q5JVL5_HUMAN
16	1805	74.5	363	2	Q6W9C6_DRYNI
17	1804	74.4	363	2	Q6W9C0_GRODE
18	1800	74.3	363	2	Q6W9C2_PEDCA
19	1799	74.3	361	2	Q6W9B8_GRODE
20	1793	74.0	363	2	Q6W9C4_GRAMU
21	1788	73.8	360	2	Q6W9D1_CASCN
22	1787	73.7	359	2	Q6W9B9_GRODE
23	1785	73.7	363	2	Q6W9B7_SPERTR
24	1785	73.7	364	2	Q6W9D2_APLRU
25	1779	73.4	362	2	Q6W9B4_ZAPPR
26	1774	73.2	364	2	Q6W9B6_TAMHU
27	1771	73.1	362	2	Q6W9B5_THOBO
28	1766	72.9	361	2	Q6W9C8_DIPSP
29	1752	72.3	359	2	Q6W9C7_DIPSA
30	1750	72.2	468	2	Q4ROX1_TETNG
31	1749	72.2	362	2	Q6W9C9_CHILA

32	1743	71.9	468	1	CNR1A_FUGRU	Q98894 fugu rubrip
33	1686	69.6	330	2	Q9BYE6_HUMAN	Q9BYE6 homo sapien
34	1686	69.6	330	2	Q9BFE4_MACMU	Q9BFE4 macaca mula
35	1686	69.6	330	2	Q9BFE5_ATEFU	Q9BFE5 atelea fusc
36	1681	69.3	330	2	Q9BFE2_CALGO	Q9BFE2 callimico g
37	1681	69.3	330	2	Q9BFE3_HYLCO	Q9BFE3 hyleobates c
38	1674	69.1	475	2	Q7T3Q3_BRAKE	Q7T3Q3 brachydanio
39	1656	68.3	330	2	Q9BFC9_TAPIN	Q9BFC9 tapirus ind
40	1656	68.3	330	2	Q9BFF6_LOXAF	Q9BFF6 loxodonta a
41	1656	68.3	330	2	Q9BFF8_TRIMA	Q9BFF8 trichechus
42	1656	68.3	330	2	Q9BFG4_TAMTE	Q9BFG4 tamandua te
43	1655	68.3	330	2	Q9BFD0_CERSI	Q9BFD0 ceratotheri
44	1653	68.2	330	2	Q9BFD3_PIG	Q9BFD3 sus scrofa
45	1652	68.2	330	2	Q8MK91_AMBHO	Q8MK91 amblysomus

#### ALIGNMENTS

RESULT 1

CNR1\_HUMAN STANDARD; PRT; 472 AA.

AC P21554; Q13949; Q9UNNO;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Cannabinoid receptor 1 (CB1) (CB-R) (CANN6).

GN Name=CNR1; Synonyms=CNR;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI\_TaxID=9606;

[1]

RP NUCLEOTIDE SEQUENCE (ISOFORM LONG).

RC TISSUE=Brain stem;

RX MEDLINE=92028798; PubMed=1718258;

RA Gerard C., Mollereau C., Vassart G., Parmentier M.;

RT "Molecular cloning of a human cannabinoid receptor which is also expressed in testis."

RL Biochem. J. 279:129-134(1991).

[2]

RP NUCLEOTIDE SEQUENCE (ISOFORM LONG).

RC TISSUE=Brain stem;

RX MEDLINE=91088303; PubMed=2263478;

RA Gerard C., Mollereau C., Vassart G., Parmentier M.;

RT "Nucleotide sequence of a human cannabinoid receptor cDNA."

RL Nucleic Acids Res. 18:7142-7142(1990).

[3]

RP NUCLEOTIDE SEQUENCE (ISOFORMS LONG AND SHORT).

RC TISSUE=Lung;

RX MEDLINE=95181329; PubMed=7876112; DOI=10.1074/jbc.270.8.3726;

RA Shire D., Carillon C., Keshad M., Calandra B., Rinaldi-Carmona M., Le Fur G., Caput D., Ferrara P.;

RT "An amino-terminal variant of the central cannabinoid receptor resulting from alternative splicing."

RL J. Biol. Chem. 270:3726-3731(1995).

[4]

RP NUCLEOTIDE SEQUENCE (ISOFORM LONG).

RC TISSUE=Hippocampus;

RX Kathmann M., Schlicker E.;

RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

[5]

RP NUCLEOTIDE SEQUENCE (ISOFORM LONG).

RA Bonner T.;

RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

[6]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM LONG).

RA Kopatz S.A., Aronstam R.S., Sharma S.V.;

RT "cDNA clones of human proteins involved in signal transduction sequenced by the Guthrie cDNA resource center (www.cdna.org).";

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

[7]

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22935763; PubMed=14574404; DOI=10.1038/nature02055;
RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
RA Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
RA Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R.,
RA Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
RA Babbage A.K., Baguley C.L., Bailey J., Banerjee R., Barker D.J.,
RA Blakesley K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
RA Blakesley S., Bray-Allen S., Brook J., Brown A.J., Brown J.V.,
RA Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
RA Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
RA Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J.,
RA Cullley K.A., Davies J., Davies J., Dunn M., Earthrowl M.E.,
RA Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
RA Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
RA Gilby L.M., Gillson C.J., Glithero R.J., Graham D.V., Grant M.,
RA Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
RA Hammond S., Harley J.B., Hart E.A., Heath P.D., Heathcote R.,
RA Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckie E.,
RA Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
RA Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
RA Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
RA Masland G.L., Matthews L., McCann O.T., McClure S.J., McKay K.,
RA McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
RA Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
RA Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
RA Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.M.,
RA Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggion L.,
RA Squares S.L., Steward C.A., Symamore N., Tamlyn-Hall G., Tester J.,
RA Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
RA Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.M.,
RA Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
RA Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
RA Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
RT "The DNA sequence and analysis of human chromosome 6."
RL Nature 425:805-811(2003).
[8]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM LONG).
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RC TISSUE=Lung;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Involved in cannabinoid-induced CNS effects. Acts by
inhibiting adenylyl cyclase. Could be a receptor for anandamide.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Names=Long;
CC IsoId=P21554-1; Sequence=Displayed;
CC Names=Short; Synonyms=CB1A;
CC IsoId=P21554-2; Sequence=VSP_001868;
CC -1- TISSUE SPECIFICITY: Found in all peripheral organs tested: heart,
colon, stomach, liver, pancreas, placenta, lung, kidney, bile
duct, spleen, adrenal gland, muscle, brain, and testis; and

```

```

CC throughout the central nervous system.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL; X54937; CAA38699.1; -; mRNA.
CC EMBL; X81120; CAA57018.1; -; mRNA.
CC EMBL; X81121; CAA57019.1; -; mRNA.
CC EMBL; AF107262; AAD34320.1; -; mRNA.
CC EMBL; U73304; AAB18200.1; -; Genomic DNA.
CC EMBL; AY225225; AAO67710.1; -; Genomic DNA.
CC EMBL; AL136096; CAB96726.1; -; Genomic DNA.
CC EMBL; BC074811; AAH74811.1; -; mRNA.
CC EMBL; BC074812; AAH74812.1; -; mRNA.
CC PIR; S17595; S17595.
CC PDB; 1LVQ; NMR; A=338-346.
CC PDB; 1LVQ; NMR; A=338-346.
CC Ensembl; ENSG00000118432; Homo sapiens.
CC HGNC; HGNC:2159; CNR1.
CC MIM; 114610; -.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0005886; C:plasma membrane; TAS.
CC GO; GO:0004949; F:cannabinoid receptor activity; TAS.
CC GO; GO:0007610; P:behavior; TAS.
CC GO; GO:0007187; P:G-protein signaling, coupled to cyclic nucl. .; TAS.
CC InterPro; IPR002230; Cnoid_receptor.
CC InterPro; IPR000810; Cnoid_receptor1.
CC InterPro; IPR000276; GPCR_Rhodopsin.
CC PANTHER; PTHR19266:SP8; Cnoid_receptor1; 1.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00552; CANNABINOIDR.
CC PRINTS; PR00362; CANNABINOIDR.
CC PRINTS; PR00237; GPCR_RHODOPSIN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
CC 3D-structure; Alternative splicing; G-protein coupled receptor;
KW Glycoprotein; Receptor; Transducer; Transmembrane.
KW TOPO_DOM 1 116 Extracellular (Potential).
FT TRANSMEM 117 142 1 (Potential).
FT TOPO_DOM 143 154 Cytoplasmic (Potential).
FT TRANSMEM 155 175 2 (Potential).
FT TOPO_DOM 176 187 Extracellular (Potential).
FT TRANSMEM 188 212 3 (Potential).
FT TOPO_DOM 213 232 Cytoplasmic (Potential).
FT TRANSMEM 233 255 4 (Potential).
FT TOPO_DOM 256 273 Extracellular (Potential).
FT TRANSMEM 274 299 5 (Potential).
FT TOPO_DOM 300 344 Cytoplasmic (Potential).
FT TRANSMEM 345 355 6 (Potential).
FT TOPO_DOM 366 377 Extracellular (Potential).
FT TRANSMEM 378 399 7 (Potential).
FT TOPO_DOM 400 472 Cytoplasmic (Potential).
FT CARBOHYD 77 77 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 83 83 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 112 112 N-linked (GlcNAc. .) (Potential).
FT VARSPLIC 1 89 MKSLDGLADTTFTTTLLVYGSNDIQYEDIKGMASKL
GYFPQKFLTSRSPFQKRWAGDNPQLVPADQVNIETFY
NKSLSF -> NALQIPPSAPSLTCTWAQMTFTSKTS
(in isoform Short).
FT FTId=VSP_001868.
FT F -> L (in Ref. 4).
FT I -> V (in Ref. 4).
FT V -> A (in Ref. 4).
CC CONFLICT 200 200
CC CONFLICT 216 216
CC CONFLICT 246 246
CC SEQUENCE 472 AA; 52858 MW; 1D2E49061D12ABF2 CRC64;
Query Match 99.5%; Score 2412; DB 1; Length 472;
Best Local Similarity 99.6%; Pred. No. 2.8e-156;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MKSILDLGLADTTTTRITITDLYVGSNDIOYEDIKGDMASKLGYFPQKPLTTSFRGSPFQE 60  
 Db 1 MKSILDLGLADTTTTRITITDLYVGSNDIOYEDIKGDMASKLGYFPQKPLTTSFRGSPFQE 60  
 QY 61 KMTAGDNPQLVPADQVNIITFYFNKSLSSFKENEENIQCGENFMDIECFMWLNFSQQLAIA 120  
 Db 61 KMTAGDNPQLVPADQVNIITFYFNKSLSSFKENEENIQCGENFMDIECFMWLNFSQQLAIA 120  
 QY 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDPHVF 180  
 Db 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDPHVF 180  
 QY 181 HRKDSRNVFLKGGVTASTASVGSIFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 Db 181 HRKDSRNVFLKGGVTASTASVGSIFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 QY 241 WTIAIVAVLPLLGWNCCKLQSVCSDFPHIDETFLMFVIGVTSVLLLFIVYAYMYILWK 300  
 Db 241 WTIAIVAVLPLLGWNCCKLQSVCSDFPHIDETFLMFVIGVTSVLLLFIVYAYMYILWK 300  
 QY 301 AHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMALIRAKTLVLVLVLIICWGPLL 360  
 Db 301 AHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMALIRAKTLVLVLVLIICWGPLL 360  
 QY 361 AIMVYDVGKMKLIKTVFAFCMLCLNSTVNPVLIYALRSKDLRHAFRSMFSPCECTAQ 420  
 Db 361 AIMVYDVGKMKLIKTVFAFCMLCLNSTVNPVLIYALRSKDLRHAFRSMFSPCECTAQ 420  
 QY 421 PLDMSMGSDCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAAEL 472  
 Db 421 PLDMSMGSDCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAAEL 472

RESULT 2  
 CNR1\_PANTR STANDARD; PRT; 472 AA.  
 ID\_QSIS73;  
 AC 10-MAY-2005 (Rel. 47, Created)  
 DT 10-MAY-2005 (Rel. 47, Last sequence update)  
 DE Cannabinoid receptor 1 (CB1) (CB-R).  
 GN Name=CNRI;  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 -OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Pan.  
 OC NCBI\_TaxID=9598;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA].  
 RX PubMed=15620360; DOI=10.1016/j.j.cell.2004.11.040;  
 RA Dorus S., Vallender E.J., Evans P.D., Anderson J.R., Gilbert S.L.,  
 RA Mahowald M., Wyckoff G.J., Malcom C.M., Lahn B.T.;  
 RT "Accelerated evolution of nervous system genes in the origin of Homo sapiens";  
 RL Cell 119:1027-1040(2004).  
 CC -!- FUNCTION: Involved in cannabinoid-induced CNS effects. Acts by inhibiting adenylate cyclase. Could be a receptor for anandamide (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
 CC -----  
 CC EMBL; AY665255; AAV74293.1; -; mRNA.  
 DR InterPro; IPR002230; Cnbdoid\_receptor.  
 DR InterPro; IPR000810; Cnoid\_receptor1.  
 DR InterPro; IPR000276; GPCR\_Rhodopsin.  
 DR PANTHER; PTHR19286:SF8; Cnoid\_receptor1; 1.  
 DR Pfam; PF00001; 7tm\_1; 1.

PRINTS; PR00522; CANNABINOIDIR.  
 PRINTS; PR00362; CANNABINOIDIR.  
 PRINTS; PR00237; GPCR\_RHODOPSIN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transducer;  
 KW Transmembrane.  
 FT TOPO\_DOM 1 116 Extracellular (Potential).  
 FT TRANSMEM 117 142 1 (Potential).  
 FT TOPO\_DOM 143 154 Cytoplasmic (Potential).  
 FT TRANSMEM 155 175 2 (Potential).  
 FT TOPO\_DOM 176 187 Extracellular (Potential).  
 FT TRANSMEM 188 212 3 (Potential).  
 FT TOPO\_DOM 213 232 Cytoplasmic (Potential).  
 FT TRANSMEM 233 255 4 (Potential).  
 FT TOPO\_DOM 256 273 Extracellular (Potential).  
 FT TRANSMEM 274 299 5 (Potential).  
 FT TOPO\_DOM 300 344 Cytoplasmic (Potential).  
 FT TRANSMEM 345 366 6 (Potential).  
 FT TOPO\_DOM 366 377 Extracellular (Potential).  
 FT TRANSMEM 378 399 7 (Potential).  
 FT TOPO\_DOM 400 472 Cytoplasmic (Potential).  
 FT CARBOHYD 77 77 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 83 83 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 472 AA; 52858 MW; 1D2E49061D12ABF2 CRC64;  
 Query Match 99.5%; Score 2412; DB 1; Length 472;  
 Best Local Similarity 99.6%; Pred. No. 2.8e-156;  
 Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MKSILDLGLADTTTTRITITDLYVGSNDIOYEDIKGDMASKLGYFPQKPLTTSFRGSPFQE 60  
 Db 1 MKSILDLGLADTTTTRITITDLYVGSNDIOYEDIKGDMASKLGYFPQKPLTTSFRGSPFQE 60  
 QY 61 KMTAGDNPQLVPADQVNIITFYFNKSLSSFKENEENIQCGENFMDIECFMWLNFSQQLAIA 120  
 Db 61 KMTAGDNPQLVPADQVNIITFYFNKSLSSFKENEENIQCGENFMDIECFMWLNFSQQLAIA 120  
 QY 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDPHVF 180  
 Db 121 VLSLTGTFVLENLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDPHVF 180  
 QY 181 HRKDSRNVFLKGGVTASTASVGSIFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 Db 181 HRKDSRNVFLKGGVTASTASVGSIFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 QY 241 WTIAIVAVLPLLGWNCCKLQSVCSDFPHIDETFLMFVIGVTSVLLLFIVYAYMYILWK 300  
 Db 241 WTIAIVAVLPLLGWNCCKLQSVCSDFPHIDETFLMFVIGVTSVLLLFIVYAYMYILWK 300  
 QY 301 AHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMALIRAKTLVLVLVLIICWGPLL 360  
 Db 301 AHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMALIRAKTLVLVLVLIICWGPLL 360  
 QY 361 AIMVYDVGKMKLIKTVFAFCMLCLNSTVNPVLIYALRSKDLRHAFRSMFSPCECTAQ 420  
 Db 361 AIMVYDVGKMKLIKTVFAFCMLCLNSTVNPVLIYALRSKDLRHAFRSMFSPCECTAQ 420  
 QY 421 PLDMSMGSDCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAAEL 472  
 Db 421 PLDMSMGSDCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAAEL 472

RESULT 3  
 Q4PLI4\_HUMAN PRELIMINARY; PRT; 472 AA.  
 ID\_Q4PLI4\_HUMAN PRELIMINARY;  
 AC Q4PLI4;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Cannabinoid receptor.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Homnidae;  
 OX Homo.  
 RN NCBI\_TaxID=9606;  
 [1]\_TaxID=9606;  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain tumor;  
 RA Kumar S., Gupta S., Shabana, Sharma, G.;  
 DR Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 CC -l- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 DR EMBL: DQ067455, AY68486.1; -, mRNA.  
 DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR InterPro: IPR002230; Cnoid receptor.  
 DR InterPro: IPR000810; Cnoid receptor.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm 1; 1.  
 DR PRINTS: PR00522; CANABINOIDIR.  
 DR PRINTS: PR00362; CANNABINOIR.  
 DR PRINTS: PR00237; GPCRHOPOPSN.  
 DR PROSITE: PS00237; G-PROTEIN RECP F1\_1; UNKNOWN 1.  
 DR PROSITE: PS0262; G-PROTEIN RECP F1\_2; 1.  
 DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.  
 KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.  
 SQ SEQUENCE 472 AA; 52858 MW; 1D2B49061D12A3F2 CRC64;  
 Query Match 99.5%; Score 2412; DB 2; Length 472;  
 Best Local Similarity 99.6%; Pred. No. 2.8e-156;  
 Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MKSILDLGLADTTFRITTTDLLVYGSNDIOYEDIKGDMASKLGVFPKFPPLTSFRGSPFOE 60  
 DB 1 MKSILDLGLADTTFRITTTDLLVYGSNDIOYEDIKGDMASKLGVFPKFPPLTSFRGSPFOE 60  
 QY 61 KMTAGNPOLVPADQVNITEFYFKSLSSPKENEENIQCGENFMDIECFVNLNPSQQLAIA 120  
 DB 61 KMTAGNPOLVPADQVNITEFYFKSLSSPKENEENIQCGENFMDIECFVNLNPSQQLAIA 120  
 QY 121 VLSLTGLTFTVLENLVLVILHRSRLCRPSYHFTGSLAVADLLGSLVIFVYSFIDFHFV 180  
 DB 121 VLSLTGLTFTVLENLVLVILHRSRLCRPSYHFTGSLAVADLLGSLVIFVYSFIDFHFV 180  
 QY 181 HRKDSRNVELFKLGGVTASFTASVGLFTATARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 DB 181 HRKDSRNVELFKLGGVTASFTASVGLFTATARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 QY 241 WTIAIVIAVLPLLGWNCCKLQVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMILWK 300  
 DB 241 WTIAIVIAVLPLLGWNCCKLQVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMILWK 300  
 QY 301 AHSNAVMIQRGTQKSIITHTSBDGKVQVTRPDQARMALRLAKTLVLILVLLICWGPLL 360  
 DB 301 AHSNAVMIQRGTQKSIITHTSBDGKVQVTRPDQARMALRLAKTLVLILVLLICWGPLL 360  
 QY 361 AIMVYDVFQGMNKLKTVFAPFCMLCLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 DB 361 AIMVYDVFQGMNKLKTVFAPFCMLCLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 QY 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 DB 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 RESULT 4  
 Q71SP5\_MACWU  
 ID Q71SP5\_MACWU PRELIMINARY; PRT; 472 AA.  
 AC Q71SP5;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Cannabinoid receptor CB-1.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;  
 OC Cercopitheciidae; Cercopitheciinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Cerebellum;  
 RA Miller G.M., Madras B.K.;  
 DR Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -l- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 DR EMBL: AF286025, AAP97250.1; -, mRNA.  
 DR GO: GO:0016021; C:integral to membrane; IEA.  
 DR GO: GO:0004949; F:cannabinoid receptor activity; IEA.  
 DR GO: GO:0004872; F:receptor activity; IEA.  
 DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
 DR GO: GO:0007165; P:signal transduction; IEA.  
 DR InterPro: IPR002230; Cnoid receptor.  
 DR InterPro: IPR000810; Cnoid receptor.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm 1; 1.  
 DR PRINTS: PR00522; CANABINOIDIR.  
 DR PRINTS: PR00362; CANNABINOIR.  
 DR PRINTS: PR00237; GPCRHOPOPSN.  
 DR PROSITE: PS00237; G-PROTEIN RECP F1\_1; UNKNOWN 1.  
 DR PROSITE: PS0262; G-PROTEIN RECP F1\_2; 1.  
 DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.  
 KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.  
 SQ SEQUENCE 472 AA; 52858 MW; 1D2E49061D12ABF2 CRC64;  
 Query Match 99.5%; Score 2412; DB 2; Length 472;  
 Best Local Similarity 99.6%; Pred. No. 2.8e-156;  
 Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MKSILDLGLADTTFRITTTDLLVYGSNDIOYEDIKGDMASKLGVFPKFPPLTSFRGSPFOE 60  
 DB 1 MKSILDLGLADTTFRITTTDLLVYGSNDIOYEDIKGDMASKLGVFPKFPPLTSFRGSPFOE 60  
 QY 61 KMTAGNPOLVPADQVNITEFYFKSLSSPKENEENIQCGENFMDIECFVNLNPSQQLAIA 120  
 DB 61 KMTAGNPOLVPADQVNITEFYFKSLSSPKENEENIQCGENFMDIECFVNLNPSQQLAIA 120  
 QY 121 VLSLTGLTFTVLENLVLVILHRSRLCRPSYHFTGSLAVADLLGSLVIFVYSFIDFHFV 180  
 DB 121 VLSLTGLTFTVLENLVLVILHRSRLCRPSYHFTGSLAVADLLGSLVIFVYSFIDFHFV 180  
 QY 181 HRKDSRNVELFKLGGVTASFTASVGLFTATARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 DB 181 HRKDSRNVELFKLGGVTASFTASVGLFTATARYISIHRLAYKRIIVTRPKAVAFCLM 240  
 QY 241 WTIAIVIAVLPLLGWNCCKLQVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMILWK 300  
 DB 241 WTIAIVIAVLPLLGWNCCKLQVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMILWK 300  
 QY 301 AHSNAVMIQRGTQKSIITHTSBDGKVQVTRPDQARMALRLAKTLVLILVLLICWGPLL 360  
 DB 301 AHSNAVMIQRGTQKSIITHTSBDGKVQVTRPDQARMALRLAKTLVLILVLLICWGPLL 360  
 QY 361 AIMVYDVFQGMNKLKTVFAPFCMLCLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 DB 361 AIMVYDVFQGMNKLKTVFAPFCMLCLNSTVNPPIIYALRSKDLRHAFRSMPPSCBGTAA 420  
 QY 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 DB 421 PLDNSMGDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
 RESULT 5  
 Q506J9\_HUMAN  
 ID Q506J9\_HUMAN PRELIMINARY; PRT; 472 AA.  
 AC Q506J9;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Cannabinoid receptor type-1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Homnidae;  
 OX Homo.  
 RN NCBI\_TaxID=9606;



```

RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Brain tumor;
RA Kumar S., Gupta S., Kariath A., Sharma G.;
RA Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
RL EMBL; AY95204; AA211179.1; -; mRNA.
DR GO; GO:0004872; F-receptor activity; IEA.
DR GO; GO:002230; G-protein coupled receptor.
DR InterPro; IPR000810; G-protein coupled receptor.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00522; CANABINOID1R.
DR PRINTS; PR00362; CANNABINOIDR.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 472 AA; 52628 MW; 49389F3C2C8CC10F CRC64;

Query Match 98.2%; Score 2380; DB 2; Length 472;
Best Local Similarity 98.5%; Pred. No. 4.3e-154;
Matches 465; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MKSLDGLADTTFTITTDLLVGSNDIQYEDIKGDMASKLGYPPQKPLTFRGSPFQE 60
Db 1 MKSLDGLADTTFTITTDLLVGSNDIQYEDIKGDMASKLGYPPQKPLTFRGSPFQE 60
Qy 61 KMTAGDNPQLVPADQVNITEFYKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 120
Db 61 KMTAGDNPQLVPADQVNITEFYKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 120
Qy 121 VLSLTGLTFTVLENLVLCVILHRSRLCRPSYHFIGSLAVADLIGSVIFVYSFIDFHFV 180
Db 121 VLSLTGLTFTVLENLVLCVILHRSRLCRPSYHFIGSLAVADLIGSVIFVYSFIDFHFV 180
Qy 181 HRKDSRNVPFLKGGVTFASVGSFLTAARYISIHRLAYKRVITRPAKVAFAFLM 240
Db 181 HRKDSRNVPFLKGGVTFASVGSFLTAARYISIHRLAYKRVITRPAKVAFAFLM 240
Qy 241 WTIAIVAVPLLGWNCCKLQVSCDIPPHIDETVLMFWIGVTSVLLFIVYAYMYLWK 300
Db 241 WTIAIVAVPLLGWNCCKLQVSCDIPPHIDETVLMFWIGVTSVLLFIVYAYMYLWK 300
Qy 301 AHSNAVMIQRTGQKSIITHTSEDKGVQVTRPDQARMAIRLAKTLVLLVLIICWGPLL 360
Db 301 AHSNAVMIQRTGQKSIITHTSEDKGVQVTRPDQARMAIRLAKTLVLLVLIICWGPLL 360
Qy 361 AIMVYDFGKMKLIKTVAFCSMLCLLNTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 420
Db 361 AIMVYDFGKMKLIKTVAFCSMLCLLNTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 420
Qy 421 PLDMSMGSDCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472
Db 421 PLDMSMGSDCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 6
Q4VB6 HUMAN PRELIMINARY; PRT; 471 AA.
ID Q4VB6 HUMAN PRELIMINARY;
AC Q4VB6;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE CNR1 protein (Fragment).
GN Name=CNR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1] NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=G-protein coupled receptors;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickens M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallus D.B.,
RA Schnarch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=G-protein coupled receptors;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC095513; AA95513.1; -; mRNA.
DR InterPro; IPR002230; Cnoid receptor.
DR InterPro; IPR000810; Cnoid receptor.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00522; CANABINOID1R.
DR PRINTS; PR00362; CANNABINOIDR.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
FT NON TER 1
FT NON TER 471
SQ SEQUENCE 471 AA; 52674 MW; 51DC852E2E43DD10 CRC64;

Query Match 98.1%; Score 2377; DB 2; Length 471;
Best Local Similarity 98.7%; Pred. No. 6.9e-154;
Matches 465; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 KSILDGLADTTFTITTDLLVGSNDIQYEDIKGDMASKLGYPPQKPLTFRGSPFQE 61
Db 1 KSILDGLADTTFTITTDLLVGSNDIQYEDIKGDMASKLGYPPQKPLTFRGSPFQE 60
Qy 62 MTAGDNPQLVPADQVNITEFYKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 121
Db 61 MTAGDNPQLVPADQVNITEFYKSLSPFKENEENIQCGENFMDIECFMVLNPSQQLAIA 120
Qy 122 LSLTLGTFTVLENLVLCVILHRSRLCRPSYHFIGSLAVADLIGSVIFVYSFIDFHFV 181
Db 121 LSLTLGTFTVLENLVLCVILHRSRLCRPSYHFIGSLAVADLIGSVIFVYSFIDFHFV 180
Qy 182 RKDSRNVPFLKGGVTFASVGSFLTAARYISIHRLAYKRVITRPAKVAFAFLM 241
Db 181 RKDSRNVPFLKGGVTFASVGSFLTAARYISIHRLAYKRVITRPAKVAFAFLM 240
Qy 242 TIAIVAVPLLGWNCCKLQVSCDIPPHIDETVLMFWIGVTSVLLFIVYAYMYLWK 301
Db 241 TIAIVAVPLLGWNCCKLQVSCDIPPHIDETVLMFWIGVTSVLLFIVYAYMYLWK 300
Qy 302 HSHAVMIQRTGQKSIITHTSEDKGVQVTRPDQARMAIRLAKTLVLLVLIICWGPLL 361
Db 301 HSHAVMIQRTGQKSIITHTSEDKGVQVTRPDQARMAIRLAKTLVLLVLIICWGPLL 360
Qy 362 IMVYDFGKMKLIKTVAFCSMLCLLNTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 421
Db 361 IMVYDFGKMKLIKTVAFCSMLCLLNTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 420

```

QY 422 LNSMGSDCLHKNNAASVHRAAESCIKSTVKIAKVTMSVSTDSAEAL 472  
 DB 421 LNSMGSDCLHKNNAASVHRAAESCIKSTVKIAKVTMSVSTDSAEAL 471

RESULT 7  
 ID CNR1\_RAT STANDARD; PRT; 473 AA.  
 AC P20272;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Cannabinoid receptor 1 (CB1) (CB-R) (Brain-type cannabinoid receptor).  
 GN Name=Cnr1; Synonyms=Skrl6;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=101116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Brain cortex;  
 RX MEDLINE=90332039; PubMed=2165569; DOI=10.1038/346561a0;  
 RA Matsuda L.A., Lolait S.J., Brownstein M.J., Young A.C., Bonner T.I.;  
 RT "Structure of a cannabinoid receptor and functional expression of the  
 RT cloned cDNA";  
 RL Nature 346:561-564 (1990).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=96429553; PubMed=8832654; DOI=10.1016/0304-3940(96)12792-0;  
 RA Ho B.Y., Zhao J.;  
 RT "Determination of the cannabinoid receptors in mouse x rat hybridoma  
 RT Ng108-15 cells and rat GH4C1 cells.";  
 RL Neurosci. Lett. 212:123-126(1996).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE OF 1-107.  
 RX MEDLINE=95181329; PubMed=7876112; DOI=10.1074/jbc.270.8.3726;  
 RA Shire D., Garillon C., Kaghad M., Calandra B., Rinaldi-Carmona M.,  
 RA Le Fur G., Caput D., Ferrara P.;  
 RT "An amino-terminal variant of the central cannabinoid receptor  
 RT resulting from alternative splicing";  
 RL J. Biol. Chem. 270:3726-3731(1995).  
 CC -1- FUNCTION: Involved in cannabinoid-induced CNS effects. Acts by  
 CC inhibiting adenylyl cyclase. Could be a receptor for anandamide.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; X55812; CA393332.1; -; mRNA.  
 DR EMBL; U40395; AAA99067.1; -; mRNA.  
 DR FIR; A33117; A33117.  
 DR Ensemble; ENSRNOG00000008223; Rattus norvegicus.  
 DR RGD; 2369; Cnr1.  
 DR InterPro; IPR002230; Cbnoid receptor.  
 DR InterPro; IPR000810; GPCR\_Rhodopsin.  
 DR InterPro; IPR000276; GPCR\_Rhodopsin.  
 DR PANTHER; PTHR19266:SF8; Cnoid\_receptor1; 1.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00522; CANNABINOIDR.  
 DR PRINTS; PR00362; CANNABINOIDR.  
 DR PRINTS; PR00237; GPCRHOOPS.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
 DR PROSITE; PS00282; G\_PROTEIN\_RECP\_F1\_2; 1.  
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transducer;  
 KW Transmembrane.  
 FT TOPO\_DOM 1 117 Extracellular (Potential).  
 FT TRANSMEM 118 143 1 (Potential).

FT TOPO\_DOM 144 155 Cytoplasmic (Potential).  
 FT TRANSMEM 156 176 2 (Potential).  
 FT TOPO\_DOM 177 188 Extracellular (Potential).  
 FT TRANSMEM 189 213 3 (Potential).  
 FT TOPO\_DOM 214 233 Cytoplasmic (Potential).  
 FT TRANSMEM 234 256 4 (Potential).  
 FT TOPO\_DOM 257 274 Extracellular (Potential).  
 FT TRANSMEM 275 300 5 (Potential).  
 FT TOPO\_DOM 301 345 Cytoplasmic (Potential).  
 FT TRANSMEM 346 366 6 (Potential).  
 FT TOPO\_DOM 367 378 Extracellular (Potential).  
 FT TRANSMEM 379 400 7 (Potential).  
 FT TOPO\_DOM 401 473 Cytoplasmic (Potential).  
 FT CARBOHYD 78 78 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 84 84 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 473 AA; 52845 MW; E59A66AFE17B646C CRC64;

Query Match 96.7%; Score 2343.5; DB 1; Length 473;  
 Best Local Similarity 96.8%; Pred. No. 1.3e-151;  
 Matches 458; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 MKSILDGLADTTFTTTDLLYVGSNDIQYEDIKGDMSKLGYPQKPLTSPRGSPPOR 60  
 DB 1 MKSILDGLADTTFTTTDLLYVGSNDIQYEDIKGDMSKLGYPQKPLTSPRGSPPOR 60

QY 61 KMTAGDNPOLVPA-DQVNITEFYNKSLSSFKENEENIOCGENFMDECFMVLNPSQOLAI 119  
 DB 61 KMTAGDNPOLVPA-DQVNITEFYNKSLSSFKENEENIOCGENFMDECFMVLNPSQOLAI 120

QY 120 AVLSLTGLTFTVLENLVLCVILHSRLCRPSYHFIGSLAVADLLGVSIFYVSFTDFHV 179  
 DB 121 AVLSLTGLTFTVLENLVLCVILHSRLCRPSYHFIGSLAVADLLGVSIFYVSFTDFHV 180

QY 180 FIRKSRNVFLKLGVTASFTASVLSLFTAIAYISIRPLAYKRIYTRKAVVAFCL 239  
 DB 181 FHRKDSPNVFLKLGVTASFTASVLSLFTAIAYISIRPLAYKRIYTRKAVVAFCL 240

QY 240 MWTIAIVAVLPLLGWCKEQLQSCSDIPPHIDEYLMFWIGVTSVLLLFIVVAYMILW 299  
 DB 241 MWTIAIVAVLPLLGWCKEQLQSCSDIPPLIDEYLMFWIGVTSVLLLFIVVAYMILW 300

QY 300 KAHSHAVRMIOQGTQKSIHHTSEDKVQVTPDQARMAIRAKTLVLVLVLIICWGPL 359  
 DB 301 KAHSHAVRMIOQGTQKSIHHTSEDKVQVTPDQARMDIRAKTLVLVLVLIICWGPL 360

QY 360 LAIMYDYVFGKMKLIKTVFAECMLCLNSTVNPPIIYALRSKDLRHAFRSPSCGTA 419  
 DB 361 LAIMYDYVFGKMKLIKTVFAECMLCLNSTVNPPIIYALRSKDLRHAFRSPSCGTA 420

QY 420 QPLDNSMGSDCLHKNNAASVHRAAESCIKSTVKIAKVTMSVSTDSAEAL 472  
 DB 421 QPLDNSMGSDCLHKNNAASVHRAAESCIKSTVKIAKVTMSVSTDSAEAL 473

RESULT 8  
 CNR1\_MOUSE  
 ID CNR1\_MOUSE STANDARD; PRT; 473 AA.  
 AC P47746;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Cannabinoid receptor 1 (CB1) (CB-R) (Brain-type cannabinoid receptor).  
 GN Name=Cnr1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=96272305; PubMed=8777318;  
 RA Chakrabarti A., Onalvi E.S., Chaudhuri G.;

RT "Cloning and sequencing of a cDNA encoding the mouse brain-type  
 RT cannabinoid receptor protein.";  
 RL DNA Seq. 5:385-388(1995).  
 RN [2]  
 RC NUCLEOTIDE SEQUENCE.  
 RA Bonner T.I.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=96429553; PubMed=8832654; DOI=10.1016/0304-3940(96)12792-0;  
 RA Ho B.Y.; Zhao J.;  
 RT "Determination of the cannabinoid receptors in mouse x rat hybridoma  
 RT NG108-15 cells and rat GH4C1 cells.";  
 RL Neurosci. Lett. 212:123-126(1996).  
 RN [4]  
 RN NUCLEOTIDE SEQUENCE.  
 RP Yuan Z.-Q.; Li L.; Qiu B.-S.; Song D.-K.;  
 RA "cDNA cloning and expression analysis of mouse cannabinoid receptor  
 RT (CB1) gene.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=129/Sv;  
 RX MEDLINE=9910599; PubMed=9888857; DOI=10.1126/science.283.5400.401;  
 RA Ledent C.; Valverde O.; Cossu G.; Petitot F.; Aubert J.F.; Beslot F.;  
 RA Boehme G.A.; Imperato A.; Pedrazzini T.; Roques B.P.; Vassart G.;  
 RA Fratta W.; Parmentier M.;  
 RT "Unresponsiveness to cannabinoids and reduced addictive effects of  
 RT opiates in CB1 receptor knockout mice.";  
 RL Science 283:401-404(1999).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Shenmen C.M.; Schuler G.D.;  
 RA Klausner R.D.; Collins P.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;  
 RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;  
 RA Hopkins R.F.; Jordan T.; Moore T.; Max S.I.; Wang J.M.; Hong L.;  
 RA Diatchenko L.; Maruina K.; Farmer A.A.; Rubin G.M.; Hong L.;  
 RA Brownstein M.J.; Ustin T.B.; Toshitsuki S.; Carninci P.; Prange C.;  
 RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullaly S.J.;  
 RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;  
 RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;  
 RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;  
 RA Fahey J.; Helton E.; Kettner M.; Madan A.; Rodriguez S.; Sanchez A.;  
 RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;  
 RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;  
 RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.;  
 RA Butterfield Y.S.N.; Krzywinski M.I.; Skalska U.; Smallos D.E.;  
 RA Schnerch A.; Schein J.E.; Jones S.J.M.; Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Involved in cannabinoid-induced CNS effects. Acts by  
 CC inhibiting adenylyl cyclase. Could be a receptor for anandamide.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; U17985; AA57202.1; -; mRNA.  
 DR EMBL; U22948; AA64413.1; -; Genomic\_DNA.  
 DR EMBL; U40709; AA91176.1; -; mRNA.  
 DR EMBL; AF153345; AAD34624.1; -; mRNA.  
 DR EMBL; Y18374; CAB42647.1; -; Genomic\_DNA.  
 DR EMBL; BC070447; AAH70447.1; -; mRNA.  
 DR EMBL; BC079564; AAH79564.1; -; mRNA.

DR Ensembl; ENSMUSG00000044288; Mus musculus.  
 DR MGI; MGI:104615; Cnrl.  
 DR GO; GO:0016021; C: integral to membrane; TAS.  
 DR InterPro; IPR002230; Cnbnoid\_receptor.  
 DR InterPro; IPR000810; Cnbnoid\_receptor.  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR PANTHER; PTHR1266:SF8; Cnoid\_receptor1; 1.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00522; CANNABINOIDIR.  
 DR PRINTS; PR00362; CANNABINOIDIR.  
 DR PRINTS; PR00237; GPCR\_Rhodopsn.  
 DR PROSITE; PS00237; G-PROTEIN\_RECPT\_F1\_1; 1.  
 DR PROSITE; PS0262; G-PROTEIN\_RECPT\_F1\_2; 1.  
 DR G-protein coupled receptor; Glycoprotein; Receptor; Transducer;  
 KW Transmembrane.  
 FT TOPO\_DOM 1 117 Extracellular (Potential).  
 FT TRANSMEM 118 143 1 (Potential).  
 FT TOPO\_DOM 144 155 Cytoplasmic (Potential).  
 FT TRANSMEM 156 176 2 (Potential).  
 FT TOPO\_DOM 177 188 Extracellular (Potential).  
 FT TRANSMEM 189 213 3 (Potential).  
 FT TOPO\_DOM 214 233 Cytoplasmic (Potential).  
 FT TRANSMEM 234 256 4 (Potential).  
 FT TOPO\_DOM 257 274 Extracellular (Potential).  
 FT TRANSMEM 275 300 5 (Potential).  
 FT TOPO\_DOM 301 345 Cytoplasmic (Potential).  
 FT TRANSMEM 346 366 6 (Potential).  
 FT TOPO\_DOM 367 378 Extracellular (Potential).  
 FT TRANSMEM 379 400 7 (Potential).  
 FT TOPO\_DOM 401 473 Cytoplasmic (Potential).  
 FT CARBOHYD 84 84 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 9 9 A -> G (in Ref. 1).  
 FT CONFLICT 115 115 S -> R (in Ref. 1).  
 FT CONFLICT 211 211 T -> R (in Ref. 1).  
 SQ SEQUENCE 473 AA; 52831 MW; E504168191CB6429 CRC64;  
 Query Match 96.6%; Score 2340.5; DB 1; Length 473;  
 Best Local Similarity 96.6%; Pred. No. 2.1e-151;  
 Matches 457; Conservative 6; Mismatches 9; Indels 1; Gaps 1;  
 Qy 1 MKSILDLGLADTTFTTTTLLYVGSNDIQVEDIKGDMASKLGYFPQKPLTSFRGSPQE 60  
 Db 1 MKSILDLGLADTTFTTTTLLYVGSNDIQVEDIKGDMASKLGYFPQKPLTSFRGSPQE 60  
 Qy 61 KMTAGDNPOLVPA-DQVNITEFYKNSLSSEKNEENIOCGNEFMDIECFMVLNPSQQLAI 119  
 Db 61 KMTAGDNPOLVPA-DQVNITEFYKNSLSSEKNEENIOCGNEFMDIECFMVLNPSQQLAI 120  
 Qy 120 AVLSLTGLTFTVLENLVLCVILHSRLRCRPSYHFGSLAVADLGLSVIFVYSPFDPHV 179  
 Db 121 AVLSLTGLTFTVLENLVLCVILHSRLRCRPSYHFGSLAVADLGLSVIFVYSPFDPHV 180  
 Qy 180 FHRKDSRNVLFLKGGVTSAGTASVGLFTTAIARYISIHRLPAYKRIIVTPKAVAFCL 239  
 Db 181 FHRKDSRNVLFLKGGVTSAGTASVGLFTTAIARYISIHRLPAYKRIIVTPKAVAFCL 240  
 Qy 240 MWTIAIVAVLPLGWNCKEKLQSCVDIPPHIDETVLMFVIGVTSVLLLFVYVYVILW 299  
 Db 241 MWTIAIVAVLPLGWNCKEKLQSCVDIPPHIDETVLMFVIGVTSVLLLFVYVYVILW 300  
 Qy 300 KAHSHAVRMIOGRQTKSIIHTSBDGKVQVTRPQARMARLAKTLVLVLVLIICWGPL 359  
 Db 301 KAHSHAVRMIOGRQTKSIIHTSBDGKVQVTRPQARMARLAKTLVLVLVLIICWGPL 360  
 Qy 360 LAIMVYDVFGQNKLIKTVFAFCNMLCLLNSTVNPPIIYALRSKDLRHAFRSMFSCGTA 419  
 Db 361 LAIMVYDVFGQNKLIKTVFAFCNMLCLLNSTVNPPIIYALRSKDLRHAFRSMFSCGTA 420  
 Qy 420 QPLDNSMGDSCLHKHANNAAVRAAESCIKSTVKIAKVTMSYSTDTSAAL 472  
 Db 421 QPLDNSMGDSCLHKHANNAAVRAAESCIKSTVKIAKVTMSYSTDTSAAL 473

```
RESULT 9
Q5SF33 MOUSE
ID Q5SF33_MOUSE PRELIMINARY; PRT; 473 AA.
AC Q5SF33;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Striatal cannabinoid receptor type 1 protein.
GN Name=Cnr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B6CBA;
RX PubMed=15606779; DOI=10.1111/j.1432-1033.2004.04460.x;
RA McCaw E.A., Hu H., Gomez G.T., Hebb A.L., Kelly M.E.,
RA Denovan-Wright E.M.;
RT "Structure, expression and regulation of the cannabinoid receptor gene
(CB1) in Huntington's disease transgenic mice.";
RL Eur. J. Biochem. 271:4909-4920(2004).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL; AY522555; AAS91801.1; -; Genomic_DNA.
DR EMBL; AY522554; AAS91800.1; -; mRNA.
DR MGI; MGI:104615; Cnr1.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR002230; Cnbnoid_receptor.
DR InterPro; IPR000810; Cnoid_receptor1.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00522; CANNABINOIDR.
DR PRINTS; PR00362; CANNABINOIDR.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.
KW SEQUENCE 473 AA; 52831 MW; E504168191C6429 CRC64;
SQ
Query Match 96.6%; Score 2340.5; DB 2; Length 473;
Best Local Similarity 96.6%; Pred. No. 2.1e-151;
Matches 457; Conservative 6; Mismatches 9; Indels 1; Gaps 1;
QY 1 MKSILDLADTTTTRTTTLLLYGSDIYVEDIKGDMASKLGYFPQKPLTSPRGSPPQE 60
DB 1 MKSILDLADTTTTRTTTLLLYGSDIYVEDIKGDMASKLGYFPQKPLTSPRGSPPQE 60
QY 61 KMTAGNPOLVPA-DQVNTTEFYNKSLSPKNEENIQCENFMDIECFMVLNPSQQLAI 119
DB 61 KMTAGNSPLVPAAGDTTNTTEFYNKSLSPKNEENIQCENFMDIECFMVLNPSQQLAI 120
QY 120 AVLSLTGLFTVLENLVLCVILHSRLCRPSYHPIGSLAVADLLGSGVIFVYSFDHFV 179
DB 121 AVLSLTGLFTVLENLVLCVILHSRLCRPSYHPIGSLAVADLLGSGVIFVYSFDHFV 180
QY 180 FHRKDSRNPLFKLGGVTSFTASVGSFLFTAJARISYIHRPLAYKRIYTRPKAVVAFCL 239
DB 181 FHRKDSFNPLFKLGGVTSFTASVGSFLFTADRYISYIHRPLAYKRIYTRPKAVVAFCL 240
QY 240 MWTIAIVLVPLLGNWCEKQSCVSDIPPHIDETVLMFWIGVTSVLLLFIVVAYVYILW 299
DB 241 MWTIAIVLVPLLGNWCKQSCVSDIPPLIDETVLMFWIGVTSVLLLFIVVAYVYILW 300
QY 300 KASHAVRMITQRGTSKSIHTSEDGKQVTRPDQARMAIRLAKTLVLVLIIICWGPL 359
DB 301 KASHAVRMITQRGTSKSIHTSEDGKQVTRPDQARMDIRLAKTLVLVLIIICWGPL 360
QY 360 LAIMVYDVGKNNKLIKTVPAFCSMCLLNSTWNPYIYALRSKDIHRFMRSPSCGTA 419
DB 361 LAIMVYDVGKNNKLIKTVPAFCSMCLLNSTWNPYIYALRSKDLHRFMRSPSCGTA 420
QY 420 QPLDMSGSDCLHKKHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472
DB 421 QPLDMSGSDCLHKKHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 473
RESULT 10
CNRL_FELCA
ID CNRL_FELCA STANDARD; PRT; 472 AA.
AC Q0277;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Cannabinoid receptor 1 (CB1) (CB-R).
GN Name=CNRL;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
OC Felinae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gebremedhin D., Lange A.R., Aebly M.R., Campbell W.B., Hillard C.J.,
RA Harder D.R.;
RT "The cannabinoid CB1 receptor is expressed in cat cerebral arterial
muscle cells and functions to inhibit L-type Ca2+ channel current.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in cannabinoid-induced CNS effects. Acts by
inhibiting adenylylate cyclase. Could be a receptor for anandamide.
CC Inhibits L-type Ca(2+) channel current.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in cerebral arterial muscle cells.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
EMBL; U94342; AAS93440.1; -; mRNA.
DR InterPro; IPR002230; Cnbnoid_receptor.
DR InterPro; IPR000810; Cnoid_receptor1.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR PANTHER; PTHR19266:SP8; Cnoid_receptor1; 1.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00522; CANNABINOIDR.
DR PRINTS; PR00362; CANNABINOIDR.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Receptor; Glycoprotein; Receptor; Transducer;
Transmembrane.
FT TOPO_DOM 1 116 Extracellular (Potential).
FT TRANSMEM 117 142 1 (Potential).
FT TOPO_DOM 143 154 Cytoplasmic (Potential).
FT TRANSMEM 155 175 2 (Potential).
FT TOPO_DOM 176 187 Extracellular (Potential).
FT TRANSMEM 188 212 3 (Potential).
FT TOPO_DOM 213 232 Cytoplasmic (Potential).
FT TRANSMEM 233 255 4 (Potential).
FT TOPO_DOM 256 273 Extracellular (Potential).
FT TRANSMEM 274 299 5 (Potential).
FT TOPO_DOM 300 344 Cytoplasmic (Potential).
FT TRANSMEM 345 365 6 (Potential).
FT TOPO_DOM 366 377 Extracellular (Potential).
FT TRANSMEM 378 399 7 (Potential).
FT TOPO_DOM 400 472 Cytoplasmic (Potential).
FT CARBOHYD 77 77 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 83 83 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 472 AA; 52887 MW; A2774DBF8F2DAF34 CRC64;
Query Match 96.3%; Score 2335; DB 1; Length 472;
```

Best Local Similarity 95.8%; Pred. No. 5e-151;		Matches 452; Conservative 13; Mismatches 7; Indels 0; Gaps 0;	
QY	1	MKSILDLGADTTTTRTTTDLVYGSNDIQYEDIKGDMASKLGYFPQKPLTSPRGSPPQE	60
Db	1	MKSILDLGADTTTTRTTTDLVYGSNDIQYEDIKGDMASKLGYFPQKPLTSPRGSPPQE	60
QY	61	KMTAGNPQLVPADQVNTTFYFNKSLSSFKENEENIQCGENFMDIECFMVLNPSQQLAIA	120
Db	61	KMTAGNSQLVPADQVNTTFYFNKSLSSYKENEENIQCGENFMDIECFMVLNPSQQLAIA	120
QY	121	VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSLVIFVYSFIDPHFV	180
Db	121	VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSLVIFVYSFVDHFV	180
QY	181	HRKDSRNVPFLKGGVTASTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFCLM	240
Db	181	HRKDSPNVPFLKGGVTASTASVGSFLTAIDRYISIHRLAYKRIIVTRPKAVAFCLM	240
QY	241	WTIAIVAVPLLGWCKEKLQSCSDIPPHIDETVLMFVIGTSLVLLFTVYAYMYILWK	300
Db	241	WTIAIVAVPLLGWCKLQSCSDIPPLIDETVLMFVIGTSLVLLFTVYAYMYILWK	300
QY	301	AHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVLVLIIICWGPLL	360
Db	301	AHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMDIRLAKTLVLVLIIICWGPLL	360
QY	361	AIMYDVFGRKMKLIKTVFAFCMCLLNSTVNPITIALRSKOLRHAFRSMFSCETAQ	420
Db	361	AIMYDVFGRKMKLIKTVFAFCMCLLNSTVNPITIALRSKOLRHAFRSMFSCETAQ	420
QY	421	PLDSMGDSCLHLKHANNAASVHRAAESCKSTVKIAKVTMSVTDTSAREAL	472
Db	421	PLDSMGDSCLHLKHANNANVHRAAENCIKNTVQIAKVTISVNTSAREAL	472
RESULT 11			
CNRI_POEGU	STANDARD;	PRT;	473 AA.
AC	P56971;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Cannabinoid receptor 1 (CBI) (CB-R).		
GN	Name=CNRI;		
OS	Poephila guttata (Zebra finch) (Taeniopygia guttata).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Passeriformes; Estrildidae;		
OC	Estrildinae; Taeniopygia.		
OX	NCBI_TaxID=59729;		
RN	(1)		
RP	NUCLEOTIDE SEQUENCE.		
RA	Soderstrom K., Johnson F.;		
RT	"Behavioral, pharmacologic and molecular characterization of a Zebra		
RT	finch CBI cannabinoid receptor.";		
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.		
CC	-!- FUNCTION: Involved in cannabinoid-induced CNS effects. Acts by		
CC	inhibiting adenylylate cyclase (By similarity).		
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.		
CC	-----		
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use as long as its content is in no way modified and this statement is not		
CC	removed.		
CC	-----		
DR	EMBL; AF255388; AAF78049.1; -; mRNA.		
DR	InterPro; IPR002230; Cnbnoid_receptor.		
DR	InterPro; IPR000810; Cnoid_receptor1.		
DR	InterPro; IPR000276; GPCR_Rhodopsn.		
DR	PANTHER; PTHR19286:SF8; Cnoid_receptor1; 1.		
DR	Pfam; PF00001; 7tm_1; 1.		

DR	PRINTS; PR00522; CANABINOIDIR.		
DR	PRINTS; PR00362; CANNABINOIDR.		
DR	PRINTS; PR00237; GPCRHHODPSN.		
DR	PROSITE; PS00227; G-PROTEIN_RECEP_F1_1; 1.		
DR	PROSITE; PS0282; G-PROTEIN_RECEP_F1_2; 1.		
KW	G-protein coupled receptor; Glycoprotein; Receptor; Transducer; Transmembrane.		
KW	Transmembrane.		
FT	TOPO_DOM 1 118 Extracellular (Potential).		
FT	TRANSMEM 119 144 1 (Potential).		
FT	TOPO_DOM 145 156 Cytoplasmic (Potential).		
FT	TRANSMEM 157 177 2 (Potential).		
FT	TOPO_DOM 178 189 Extracellular (Potential).		
FT	TRANSMEM 190 214 3 (Potential).		
FT	TOPO_DOM 215 234 Cytoplasmic (Potential).		
FT	TRANSMEM 235 257 4 (Potential).		
FT	TOPO_DOM 258 275 Extracellular (Potential).		
FT	TRANSMEM 276 301 5 (Potential).		
FT	TOPO_DOM 302 346 Cytoplasmic (Potential).		
FT	TRANSMEM 347 367 6 (Potential).		
FT	TOPO_DOM 368 379 Extracellular (Potential).		
FT	TRANSMEM 380 401 7 (Potential).		
FT	TOPO_DOM 402 473 Cytoplasmic (Potential).		
FT	CARBOHYD 79 79 N-linked (GlcNAc...) (Potential).		
FT	CARBOHYD 85 85 N-linked (GlcNAc...) (Potential).		
SQ	SEQUENCE 473 AA; 96EC6D1BD25461E CRC64;		
Query Match 92.8%; Score 2250.5; DB 1; Length 473;			
Best Local Similarity 91.4%; Pred. No. 2.9e-145;			
Matches 433; Conservative 25; Mismatches 13; Indels 3; Gaps 2;			
QY	1	MKSILDLGADTTTTRTTTDLVYGSNDIQYEDIKGDMASKLGYFPQKPLTSPRGSPPQE	60
Db	1	MKSILDLGADTTTTRTTTDLVYGSNDIQYEDIKGDMASKLGYFPQKPLTSPRGSPPQE	60
QY	61	KMTAGNP--QLVPADQVNTTFYFNKSLSSFKENEENIQCGENFMDIECFMVLNPSQQLA	118
Db	61	KMTGGDLSLSIIIPSEQVNITEFYFNKSLSTFKONEENIQCGENFMDIECFMVLNPSQQLA	120
QY	119	IAVLSLTGTFVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSLVIFVYSFIDPH	178
Db	121	IAVLSLTGTFVLENLVLCVILHSRSLRCRPSYHFHIGSLAVADLLGSLVIFVYSFVDHF	180
QY	179	VFHRKDSRNVPFLKGGVTASTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVAFVC	238
Db	181	VFHRKDSRNVPFLKGGVTASTASVGSFLTAIDRYISIHRLAYKRIIVTRPKAVAFVC	240
QY	239	LMWTIAIVAVPLLGWCKEKLQSCSDIPPHIDETVLMFVIGTSLVLLFTVYAYMYIL	298
Db	241	VMWTIAIVAVPLLGWCKLNSVCSDIPLIDETVLMFVIGTSLVLLFTVYAYMYIL	300
QY	299	WKAHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVLVLIIICWGVP	358
Db	301	WKAHSHAVRMIOGTOKSIIHTSEDKGVQVTRPDQARMDIRLAKTLVLVLIIICWGVP	360
QY	359	LLAIVYDVFGRKMKLIKTVFAFCMCLLNSTVNPITIALRSKOLRHAFRSMFSCBGT	418
Db	361	LLAIVYDVFGRKMKLIKTVFAFCMCLLNSTVNPITIALRSKOLRHAFRSMFSCBGT	420
QY	419	AQPLDSMGDSCLHLKHANNAASVHRAAESCKSTVKIAKVTMSVTDTSAREAL	472
Db	421	AQPLDSMGDSCLHLKHANNAGNVHRAAESCKSTVKIAKVTMSVTDTSAREAL	473
RESULT 12			
Q5UB37_HUMAN	PRELIMINARY;	PRT;	439 AA.
ID	Q5UB37_HUMAN		
AC	Q5UB37;		
DT	01-FEB-2005 (TrEMBLrel. 29, Created)		
DT	01-FEB-2005 (TrEMBLrel. 29, Last sequence update)		
DE	01-FEB-2005 (TrEMBLrel. 29, Last annotation update)		
DE	Cannabinoid receptor 1 splice variant CBIb.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX PubMed=15620723; DOI=10.1016/j.febslet.2004.11.085;  
 RA Ryberg E., Vu H.K., Larsson N., Groblewski T., Hjorth S., Elebring T.,  
 RA Sjogren S., Greasley P.J.;  
 RT "Identification and characterisation of a novel splice variant of the  
 RT human CB1 receptor.";  
 RL FEBS Lett. 579:259-264(2005).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Khang Vu H., Groblewski T., Greasley P.;  
 RA Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.  
 DR EMBL; AY766182; AAV3030.1; -, mRNA.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004949; F:cannabinoid receptor activity; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR002230; Cnbnoid receptor.  
 DR InterPro; IPR000810; Cnoid receptor1.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00522; CANABINOIDIR.  
 DR PRINTS; PR00362; CANABINOIDR.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.  
 SQ SEQUENCE 439 AA; 43060 MW; C80C6B8D640412C1 CRC64;

Query Match 91.2%; Score 2209.5; DB 2; Length 439;  
 Best Local Similarity 92.6%; Pred. No. 1.7e-142;  
 Matches 437; Conservative 0; Mismatches 12; Indels 33; Gaps 1;  
 QY 1 MXSILDLGLADTTFRITTTDLLVYGSNDIQVEDIKGDMASKLGYFPQKFLTSFRGSPFQ 60  
 DB 1 MXSILDLGLADTTFRITTTDLLVYGSNDIQVEDIKGDMASKLGYFPQKFLTSFRGSPFQ 27  
 QY 61 KMTAGNPNQLVPADQVNIETFYNKSLSFKENEENIQCGENFMDFIECFMVLNPSQQLA 120  
 DB 28 KMTAGNPNQLVPADQVNIETFYNKSLSFKENEENIQCGENFMDFIECFMVLNPSQQLA 87  
 QY 121 VLSLTIGTFTVLENLVLVILHSRLCRPSVHFHFGSLAVADLLGSLVFVYSFIDFHFV 180  
 DB 88 VLSLTIGTFTVLENLVLVILHSRLCRPSVHFHFGSLAVADLLGSLVFVYSFIDFHFV 147  
 QY 181 HRKDSNRVLFKLGVTASTASVGSIFLTAIRYISIHRLAYKRIIVTRPKAVAFCLM 240  
 DB 148 HRKDSNRVLFKLGVTASTASVGSIFLTAIRYISIHRLAYKRIIVTRPKAVAFCLM 207  
 QY 241 WTIAIVIAVLPLGWNCEKLSQVCSDFPHIDTYLMFMIGVTSVLLLFVYAYMILWK 300  
 DB 208 WTIAIVIAVLPLGWNCEKLSQVCSDFPHIDTYLMFMIGVTSVLLLFVYAYMILWK 267  
 QY 301 AHSNAVMTQRTQKSIHHTSDGKQVTRPQQAARLAKTLVILVLLIICWGPLL 360  
 DB 268 AHSNAVMTQRTQKSIHHTSDGKQVTRPQQAARLAKTLVILVLLIICWGPLL 327  
 QY 361 AIMVYDFGQNKLIKTVFAFCMCLLNSTNPILYALRSKDLRAFRMPSPSCGTAQ 420  
 DB 328 AIMVYDFGQNKLIKTVFAFCMCLLNSTNPILYALRSKDLRAFRMPSPSCGTAQ 387  
 QY 421 PLDNSMGDSCLHKKHANNAAHVRAAESCIKSTVKIAKVTMSVTDTSAAEL 472  
 DB 388 PLDNSMGDSCLHKKHANNAAHVRAAESCIKSTVKIAKVTMSVTDTSAAEL 439

RESULT 13

CNRI\_TARGR STANDARD; PRT; 473 AA.  
 ID CNRI\_TARGR STANDARD; PRT; 473 AA.  
 AC Q9PUI7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Cannabinoid receptor 1 (CB1) (CB-R).  
 GN Name=CNRI;  
 OS Taricha granulosa (Roughskin newt).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Taricha.  
 OX NCBI\_TaxID=8321;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20312895; PubMed=10854287;  
 RA Soderstrom K., Leid M., Moore F.L., Murray T.F.;  
 RT "Behavioral, pharmacological, and molecular characterization of an  
 RT amphibian cannabinoid receptor.";  
 RL J. Neurochem. 75:413-423(2000).  
 CC -1- FUNCTION: Involved in cannabinoid-induced CNS effects. Acts by  
 CC inhibiting adenylylate cyclase (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; AF181894; AAD56029.1; -, mRNA.  
 DR InterPro; IPR002230; Cnbnoid receptor.  
 DR InterPro; IPR000810; Cnoid receptor1.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR PANTHER; PTHR19266:SP8; Cnoid\_receptor1; 1.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00522; CANABINOIDIR.  
 DR PRINTS; PR00362; CANABINOIDR.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transducer;  
 KW Transmembrane.  
 FT TOPO\_DOM 1 118 Extracellular (Potential).  
 FT TRANSMEM 119 144 1 (Potential).  
 FT TOPO\_DOM 145 156 Cytoplasmic (Potential).  
 FT TRANSMEM 157 177 2 (Potential).  
 FT TOPO\_DOM 178 199 Extracellular (Potential).  
 FT TRANSMEM 190 214 3 (Potential).  
 FT TOPO\_DOM 215 234 Cytoplasmic (Potential).  
 FT TRANSMEM 235 257 4 (Potential).  
 FT TOPO\_DOM 258 275 Extracellular (Potential).  
 FT TRANSMEM 276 301 5 (Potential).  
 FT TOPO\_DOM 302 346 Cytoplasmic (Potential).  
 FT TRANSMEM 347 367 6 (Potential).  
 FT TOPO\_DOM 368 379 Extracellular (Potential).  
 FT TRANSMEM 380 401 7 (Potential).  
 FT TOPO\_DOM 402 473 Cytoplasmic (Potential).  
 FT CARBOHYD 79 79 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 85 85 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 473 AA; 53325 MW; F17C016516ED4724 CRC64;  
 Query Match 85.2%; Score 2064.5; DB 1; Length 473;  
 Best Local Similarity 83.5%; Pred. No. 1.4e-132;  
 Matches 396; Conservative 39; Mismatches 36; Indels 3; Gaps 2;  
 QY 1 MXSILDLGLADTTFRITTTDLLVYGSNDIQVEDIKGDMASKLGYFPQKFLTSFRGSPFQ 60  
 DB 1 MXSILDLGLADTTFRITTTDLLVYGSNDIQVEDIKGDMASKLGYFPQKFLTSFRGSPFQ 60  
 QY 61 KMTAGNPNQL--VPADQVNIETFYNKSLSFKENEENIQCGENFMDFIECFMVLNPSQQLA 118  
 DB 61 KMTIGDNNLLSFYPLDQFNVTFFNRSVSTFFKENDNLCGENFMDFMECFMILTASQQLI 120





KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.  
SQ SEQUENCE 411 AA; 45874 MW; E3C31ACAB4066BC1 CRC64;

Query Match		80.6%;	Score 1954;	DB 2;	Length 411;
Best Local Similarity		98.5%;	Pred. No. 4.le-125;		
Matches 382;		Conservative 1;	Mismatches 5;	Indels 0;	Gaps 0;
QY	85	SASSFKENBENIQCGNFMDIECFMVLNPSQQLAIAVLSLTGLTFTVLENLLVLCVILHS	144		
Db	24	SYKTSKENEENIQCGNFMDIECFMVLNPSQQLAIAVLSLTGLTFTVLENLLVLCVILHS	83		
QY	145	RSLRCRPSYHFIGSLAVADLLGSVIFVYGFIDPHVFRKDSRNVLFLKLGVTASTASV	204		
Db	84	RSLRCRPSYHFIGSLAVADLLGSVIFVYGFIDPHVFRKDSRNVLFLKLGVTASTASV	143		
QY	205	GSLFLTAIARYISIHREPLAYKRIVTRPKAVAFCLMWTIAIVIAVLPLLGWNCCKLOSVC	264		
Db	144	GSLFLTAIDRYISIHREPLAYKRIVTRPKAVAFCLMWTIAIVIAVLPLLGWNCCKLOSVC	203		
QY	265	SDIFFPHIDETYLMTFVIGVTSVLLLFIVYAYMYILWKASHAVRMIOQTQKSIIIHTSED	324		
Db	204	SDIFFPHIDETYLMTFVIGVTSVLLLFIVYAYMYILWKASHAVRMIOQTQKSIIIHTSED	263		
QY	325	GKQVTRPDQARMIRLAKTLVLILVLLIICWGPLLAIMVYDVFGKNNKLIKTVFAPCSM	384		
Db	264	GKQVTRPDQARMIRLAKTLVLILVLLIICWGPLLAIMVYDVFGKNNKLIKTVFAPCSM	323		
QY	385	LCILNSTVNPPIIYALRSKDLRHAFRSMFPSCGTAQPLDNSMGDSDCILHKHANNAAVHR	444		
Db	324	LCILNSTVNPPIIYALRSKDLRHAFRSMFPSCGTAQPLDNSMGDSDCILHKHANNAAVHR	383		
QY	445	AAESCICKSTVKIAKVTMSVSTDTSAEAL	472		
Db	384	AAESCICKSTVKIAKVTMSVSTDTSAEAL	411		

Search completed: January 6, 2006, 21:12:26  
Job time : 235 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 21:05:18 ; Search time 167 Seconds  
(without alignments)

1180.931 Million cell updates/sec

Title: US-10-521-420-1

Perfect score: 2424

Sequence: 1 MKSILDGLADTTTFTTTDL.....TVKIAKVTMSVSTDSAEAL 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2424	100.0	472	5	US-10-521-420-1
2	2412	99.5	472	4	US-10-225-567A-84
3	2412	99.5	472	4	US-10-029-386-33685
4	2412	99.5	472	4	US-10-295-027-236
5	2412	99.5	472	4	US-10-764-425-168
6	2412	99.5	472	5	US-10-482-029-172
7	2412	99.5	472	3	US-11-019-829-88
8	2407	99.3	472	3	US-09-826-509-469
9	2407	99.3	472	5	US-10-925-095-469
10	2401	99.1	472	4	US-10-295-027-238
11	1954	80.6	411	4	US-10-295-027-240
12	1954	80.6	411	4	US-10-408-765A-2310
13	1954	80.6	411	6	US-11-019-829-89
14	710	29.3	360	4	US-10-225-567A-86
15	710	29.3	360	4	US-10-290-078-3
16	710	29.3	360	5	US-10-473-127-1541
17	710	29.3	360	5	US-10-473-127-1542
18	710	29.3	360	5	US-10-473-127-1543
19	710	29.3	360	5	US-10-473-127-1544
20	705	29.1	360	3	US-09-826-509-471
21	705	29.1	360	5	US-10-925-095-471
22	410	16.9	80	3	US-09-993-844-50
23	410	16.9	80	4	US-10-633-438-36
24	410	16.9	80	5	US-10-693-164-38
25	410	16.9	80	5	US-10-901-772-36
26	410	16.9	80	6	US-11-026-435-50
27	369.5	15.2	362	3	US-09-993-844-8

28	369.5	15.2	362	4	US-10-633-438-65	Sequence 65, Appl
29	369.5	15.2	362	5	US-10-901-772-65	Sequence 65, Appl
30	369.5	15.2	362	6	US-11-026-435-8	Sequence 8, Appl
31	366	15.1	391	3	US-09-904-099-3	Sequence 3, Appl
32	364.5	15.0	384	3	US-09-904-099-4	Sequence 4, Appl
33	364	15.0	364	3	US-09-904-099-30	Sequence 30, Appl
34	364	15.0	364	4	US-10-228-762-15	Sequence 15, Appl
35	364	15.0	364	4	US-10-648-593-166	Sequence 166, App
36	364	15.0	364	4	US-10-764-425-130	Sequence 130, App
37	364	15.0	364	4	US-10-741-601-489	Sequence 489, App
38	364	15.0	364	4	US-10-741-601-490	Sequence 490, App
39	364	15.0	364	4	US-10-741-601-491	Sequence 491, App
40	364	15.0	364	4	US-10-741-601-492	Sequence 492, App
41	364	15.0	364	4	US-10-741-601-493	Sequence 493, App
42	364	15.0	364	5	US-10-741-600-1409	Sequence 1409, Ap
43	364	15.0	364	5	US-10-741-600-1410	Sequence 1410, Ap
44	364	15.0	364	5	US-10-741-600-1411	Sequence 1411, Ap
45	364	15.0	364	5	US-10-741-600-1412	Sequence 1412, Ap

#### ALIGNMENTS

##### RESULT 1

US-10-521-420-1

; Sequence 1, Application US/10521420

; Publication No. US20050239133A1

; GENERAL INFORMATION:

; APPLICANT: Greasley, Peter

; TITLE OF INVENTION: Methods to Identify True Antagonists and Inverse Agonists of the

; FILE REFERENCE: ASZD-P01-752

; CURRENT APPLICATION NUMBER: US/10/521.420

; CURRENT FILING DATE: 2005-01-14

; PRIOR APPLICATION NUMBER: SE 0202242-4

; PRIOR FILING DATE: 2002-07-17

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 472

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-521-420-1

Query Match	100.0%	Score 2424;	DB 5;	Length 472;
Best Local Similarity	100.0%	Pred. No. 5.3e-215;		
Matches 472;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKSILDGLADTTTFTTTTDLVYGSNDIQYEDIKGDMASKLGYFPKFPPLTSFRGSPFQE	60	
Db	1	MKSILDGLADTTTFTTTTDLVYGSNDIQYEDIKGDMASKLGYFPKFPPLTSFRGSPFQE	60	
Qy	61	KMTAGDNPQLVPADQVNITEFYNKSLSFKNENIQCENFMNDIECFMVLNPSQOLAIA	120	
Db	61	KMTAGDNPQLVPADQVNITEFYNKSLSFKNENIQCENFMNDIECFMVLNPSQOLAIA	120	
Qy	121	VLSITIGTFTVLENLVLCVILHSRSRCPSVHFICGLAVADLLGSVIFYVSIIDHFVF	180	
Db	121	VLSITIGTFTVLENLVLCVILHSRSRCPSVHFICGLAVADLLGSVIFYVSIIDHFVF	180	
Qy	181	HRKDSRNVFVFLKGGVTASFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVVAFCLM	240	
Db	181	HRKDSRNVFVFLKGGVTASFTASVGSFLTAIARYISIHRLAYKRIIVTRPKAVVAFCLM	240	
Qy	241	WTIAIVAVLPVLLGNCKLQSCSDIFPHIDEYLFMFWIGVTSVLLLFIVYAYMYILWK	300	
Db	241	WTIAIVAVLPVLLGNCKLQSCSDIFPHIDEYLFMFWIGVTSVLLLFIVYAYMYILWK	300	
Qy	301	AHSHAVRMIOQGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVILVLLVLLICWGP	360	
Db	301	AHSHAVRMIOQGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVILVLLVLLICWGP	360	
Qy	361	AIMVYDVGKMKLIKTVFAFCMCLLNSTVNPVIALRSKDLRHAFRSMFPSCGTAQ	420	

```

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEONICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33685
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL136096.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.22
; OTHER INFORMATION: SWISSPROT HIT: P21554, EVALU = 0.00e+00
US-10-029-386-33685

Query Match 99.5%; Score 2412; DB 4; Length 472;
Best Local Similarity 99.6%; Pred. No. 6.8e-214;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKSILDGLADTTFRITTTDLLLVGSDIOYEDIKGMASKLGYFPQKPLTSFRGSPFOE 60
Db 1 MKSILDGLADTTFRITTTDLLLVGSDIOYEDIKGMASKLGYFPQKPLTSFRGSPFOE 60
QY 61 KMTAGNPQLVPADQVNITEFYNKSLSFKENEENIQGENFMDIECFWLNPSQOLAIA 120
Db 61 KMTAGNPQLVPADQVNITEFYNKSLSFKENEENIQGENFMDIECFWLNPSQOLAIA 120
QY 121 VLSLTGTFVLENLLVLCVILHSRLRCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180
Db 121 VLSLTGTFVLENLLVLCVILHSRLRCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180
QY 181 HRKDSRNVLFLKGGVTASFTASVGSFLTAIARIYSIHRPLAYKRIIVTRPKAVAFCLM 240
Db 181 HRKDSRNVLFLKGGVTASFTASVGSFLTAIARIYSIHRPLAYKRIIVTRPKAVAFCLM 240
QY 241 WTIAIVAVLPILGNCKEQLQSCSDIFPHIDETVLMFWIGVTSVLLLFIVVAYMYLWK 300
Db 241 WTIAIVAVLPILGNCKEQLQSCSDIFPHIDETVLMFWIGVTSVLLLFIVVAYMYLWK 300
QY 301 AHSHAVRMIQRTQKSIITHTSDEGKVQVTRPDQARMALRLAKTLVLVLVLIICWGPPL 360
Db 301 AHSHAVRMIQRTQKSIITHTSDEGKVQVTRPDQARMALRLAKTLVLVLVLIICWGPPL 360
QY 361 AIMVYDVFQGNKLIKTVPAFCSMCLLNSTVNPITYALRSKDLRHAFRSMFPCSGTAQ 420
Db 361 AIMVYDVFQGNKLIKTVPAFCSMCLLNSTVNPITYALRSKDLRHAFRSMFPCSGTAQ 420
QY 421 PLDSNGSDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472
Db 421 PLDSNGSDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 4
US-10-295-027-236
; Sequence 236, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
```

```

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEONICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33685
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL136096.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.22
; OTHER INFORMATION: SWISSPROT HIT: P21554, EVALU = 0.00e+00
US-10-029-386-33685

Query Match 99.5%; Score 2412; DB 4; Length 472;
Best Local Similarity 99.6%; Pred. No. 6.8e-214;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKSILDGLADTTFRITTTDLLLVGSDIOYEDIKGMASKLGYFPQKPLTSFRGSPFOE 60
Db 1 MKSILDGLADTTFRITTTDLLLVGSDIOYEDIKGMASKLGYFPQKPLTSFRGSPFOE 60
QY 61 KMTAGNPQLVPADQVNITEFYNKSLSFKENEENIQGENFMDIECFWLNPSQOLAIA 120
Db 61 KMTAGNPQLVPADQVNITEFYNKSLSFKENEENIQGENFMDIECFWLNPSQOLAIA 120
QY 121 VLSLTGTFVLENLLVLCVILHSRLRCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180
Db 121 VLSLTGTFVLENLLVLCVILHSRLRCRPSYHFHIGSLAVADLLGSLVFVYSFIDFHFV 180
QY 181 HRKDSRNVLFLKGGVTASFTASVGSFLTAIARIYSIHRPLAYKRIIVTRPKAVAFCLM 240
Db 181 HRKDSRNVLFLKGGVTASFTASVGSFLTAIARIYSIHRPLAYKRIIVTRPKAVAFCLM 240
QY 241 WTIAIVAVLPILGNCKEQLQSCSDIFPHIDETVLMFWIGVTSVLLLFIVVAYMYLWK 300
Db 241 WTIAIVAVLPILGNCKEQLQSCSDIFPHIDETVLMFWIGVTSVLLLFIVVAYMYLWK 300
QY 301 AHSHAVRMIQRTQKSIITHTSDEGKVQVTRPDQARMALRLAKTLVLVLVLIICWGPPL 360
Db 301 AHSHAVRMIQRTQKSIITHTSDEGKVQVTRPDQARMALRLAKTLVLVLVLIICWGPPL 360
QY 361 AIMVYDVFQGNKLIKTVPAFCSMCLLNSTVNPITYALRSKDLRHAFRSMFPCSGTAQ 420
Db 361 AIMVYDVFQGNKLIKTVPAFCSMCLLNSTVNPITYALRSKDLRHAFRSMFPCSGTAQ 420
QY 421 PLDSNGSDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472
Db 421 PLDSNGSDSCLHGHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 3
US-10-029-386-33685
; Sequence 33685, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
```

```
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 236
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-236

Query Match      99.5%; Score 2412; DB 4; Length 472;
Best Local Similarity 99.6%; Pred. No. 6.8e-214;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKSILDGLADTTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPKFPPLTSFRGSPFQE 60
Db 1 MKSILDGLADTTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPKFPPLTSFRGSPFQE 60

Qy 61 KMTAGDNPOLVPADQVNITEFYKSLSSFKENENIOCGENFMDIECFWLNPSQQLAIA 120
Db 61 KMTAGDNPOLVPADQVNITEFYKSLSSFKENENIOCGENFMDIECFWLNPSQQLAIA 120

Qy 121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFYYSFIDFHFV 180
Db 121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFYYSFIDFHFV 180

Qy 181 HRKDSRNVPFLKGGVTASFTASVGSLSFLTAIRYISIHRLAYKRIVTRPKAVVAFCLM 240
Db 181 HRKDSRNVPFLKGGVTASFTASVGSLSFLTAIRYISIHRLAYKRIVTRPKAVVAFCLM 240

Qy 241 WTAIVAVLPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMYILWK 300
Db 241 WTAIVAVLPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMYILWK 300

Qy 301 AHSHAVRMIQRGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVLVLVLIICWGPILL 360
Db 301 AHSHAVRMIQRGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVLVLVLIICWGPILL 360

Qy 361 AIMVYDVFQKMKLIKTVFAPCSMLCLLNSTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 420
Db 361 AIMVYDVFQKMKLIKTVFAPCSMLCLLNSTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 420

Qy 421 PLDSNMGSDCLHGHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472
Db 421 PLDSNMGSDCLHGHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 5
US-10-764-425-168
; Sequence 168, Application US/10764425

; Publication No. US20040146921A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; APPLICANT: Taylor, Ian
; TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE
; FILE REFERENCE: 5151
; CURRENT APPLICATION NUMBER: US/10/764,425
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: 60/442,582
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 168
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-764-425-168

Query Match      99.5%; Score 2412; DB 4; Length 472;
Best Local Similarity 99.6%; Pred. No. 6.8e-214;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKSILDGLADTTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPKFPPLTSFRGSPFQE 60
Db 1 MKSILDGLADTTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPKFPPLTSFRGSPFQE 60

Qy 61 KMTAGDNPOLVPADQVNITEFYKSLSSFKENENIOCGENFMDIECFWLNPSQQLAIA 120
Db 61 KMTAGDNPOLVPADQVNITEFYKSLSSFKENENIOCGENFMDIECFWLNPSQQLAIA 120

Qy 121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFYYSFIDFHFV 180
Db 121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFYYSFIDFHFV 180

Qy 181 HRKDSRNVPFLKGGVTASFTASVGSLSFLTAIRYISIHRLAYKRIVTRPKAVVAFCLM 240
Db 181 HRKDSRNVPFLKGGVTASFTASVGSLSFLTAIRYISIHRLAYKRIVTRPKAVVAFCLM 240

Qy 241 WTAIVAVLPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMYILWK 300
Db 241 WTAIVAVLPLLGWNCCKLQSVCSDFPHIDETVLMFWIGVTSVLLLFIVVAYMYILWK 300

Qy 301 AHSHAVRMIQRGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVLVLVLIICWGPILL 360
Db 301 AHSHAVRMIQRGTOKSIIHTSEDKGVQVTRPDQARMAIRLAKTLVLVLVLIICWGPILL 360

Qy 361 AIMVYDVFQKMKLIKTVFAPCSMLCLLNSTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 420
Db 361 AIMVYDVFQKMKLIKTVFAPCSMLCLLNSTVNPPIIYALRSKDLRHAFRSPFSCGTAQ 420

Qy 421 PLDSNMGSDCLHGHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472
Db 421 PLDSNMGSDCLHGHANNAASVHRAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 6
US-10-482-029-172
; Sequence 172, Application US/10482029
; Publication No. US20050037445A1
; GENERAL INFORMATION:
; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482,029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 172
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
```



301 AHSHAVRMIOGTOKSIIHTSEDKVQVTRPDQARMAIRLAKTLVLVLIIICWGPLL 360  
301 AHSHAVRMIOGTOKSIIHTSEDKVQVTRPDQARMDIRLKTTLVLVLIIICWGPLL 360  
361 AIMYDVFGKONKLIKTVFAFCMLCLLNSTVNPPIIYALRSKDLRHAFRSMFSPCEGTAQ 420  
361 AIMYDVFGKONKLIKTVFAFCMLCLLNSTVNPPIIYALRSKDLRHAFRSMFSPCEGTAQ 420  
421 PLDSMGSDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
421 PLDSMGSDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 9  
US-10-925-095-469  
; Sequence 469, Application US/10925095  
; Publication No. US20050019840A1  
; GENERAL INFORMATION:  
; APPLICANT: Lehmann-Bruinsma, Karin  
; APPLICANT: Liaw, Chen W.  
; APPLICANT: Lin, I-Lin  
; TITLE OF INVENTION: Non-Endogenous, Constitutively Activated Known G  
; TITLE OF INVENTION: Protein-Coupled Receptors  
; FILE REFERENCE: AREN-207  
; CURRENT APPLICATION NUMBER: US/10/925,095  
; CURRENT FILING DATE: 2004-08-24  
; PRIOR APPLICATION NUMBER: US/09/826,509  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: 60/195,747  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: 09/170,496  
; PRIOR FILING DATE: 1998-10-13  
; NUMBER OF SEQ ID NOS: 589  
; SOFTWARE: Patent In Version 2.1  
; SEQ ID NO 469  
; LENGTH: 472  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-925-095-469

Query Match 99.3%; Score 2407; DB 5; Length 472;  
Best Local Similarity 99.4%; Pred. No. 2e-213;  
Matches 469; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 MKSILDGLADTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPQKPLTSFRGSPQE 60  
1 MKSILDGLADTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPQKPLTSFRGSPQE 60  
61 KMTAGDNPOLPADQVNITEFYNKSLSFKENEIQCENFMDECFWLNPSQOLAIA 120  
61 KMTAGDNPOLPADQVNITEFYNKSLSFKENEIQCENFMDECFWLNPSQOLAIA 120  
121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFVYSFIDFHFV 180  
121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFVYSFIDFHFV 180  
181 HRKDSRNVPFLKLGVTASFTASVGSFLTAIARVISIHRPLAYKRIIVTRPKAVVAFCLM 240  
181 HRKDSRNVPFLKLGVTASFTASVGSFLTAIARVISIHRPLAYKRIIVTRPKAVVAFCLM 240  
241 WTIAIVIAVLPILGNCEKLSQVCSDDIPPHIDETVLMFWIGVTSVLLFIIVYAVMYILWK 300  
241 WTIAIVIAVLPILGNCEKLSQVCSDDIPPHIDETVLMFWIGVTSVLLFIIVYAVMYILWK 300  
301 AHSHAVRMIOGTOKSIIHTSEDKVQVTRPDQARMAIRLAKTLVLVLIIICWGPLL 360  
301 AHSHAVRMIOGTOKSIIHTSEDKVQVTRPDQARMDIRLKTTLVLVLIIICWGPLL 360  
361 AIMYDVFGKONKLIKTVFAFCMLCLLNSTVNPPIIYALRSKDLRHAFRSMFSPCEGTAQ 420  
361 AIMYDVFGKONKLIKTVFAFCMLCLLNSTVNPPIIYALRSKDLRHAFRSMFSPCEGTAQ 420

421 PLDSMGSDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472  
421 PLDSMGSDSCLHKKHANNAASVHRAAESCIKSTVKIAKVTMSVSTDTSAEAL 472

RESULT 10  
US-10-295-027-238  
; Sequence 238, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 238  
; LENGTH: 472  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-238

Query Match 99.1%; Score 2401; DB 4; Length 472;  
Best Local Similarity 98.9%; Pred. No. 7.1e-213;  
Matches 467; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

1 MKSILDGLADTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPQKPLTSFRGSPQE 60  
1 MKSILDGLADTTFTTTTDLVYGSNDIQYEDIKGMASKLGYFPQKPLTSFRGSPQE 60  
61 KMTAGDNPOLPADQVNITEFYNKSLSFKENEIQCENFMDECFWLNPSQOLAIA 120  
61 KMTAGDNPOLPADQVNITEFYNKSLSFKENEIQCENFMDECFWLNPSQOLAIA 120  
121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFVYSFIDFHFV 180  
121 VLSLTGTFVLENLVLCVILHSRSLRCRPSYHFIGSLAVADLLGSVIFVYSFIDFHFV 180  
181 HRKDSRNVPFLKLGVTASFTASVGSFLTAIARVISIHRPLAYKRIIVTRPKAVVAFCLM 240  
181 HRKDSRNVPFLKLGVTASFTASVGSFLTAIARVISIHRPLAYKRIIVTRPKAVVAFCLM 240









Qy	138	LCVILHSRSLRCPYSVHFISGLAVADLIGSVIFVYSFIDFHVHFKDSRNVFLFKLGGVT	197
Db	55	LYLILSSHQURRKPSYLFISGLAGDFLASVFPACSFVNFHVHFGVDSKAVFLLKIGSVT	114
Qy	198	ASFTASVGSFLFTAIARYISIHRLAYKRIVTRPKAVAFCLMWTIAIVIAVPLPLGMNC	257
Db	115	MTFTASVGSILLTAIDRYLCRLYPSPYKALLTRGRALVTLGIMWVLSALVSVLPLMGWTC	174
Qy	258	EKLQSVCSDFPHIDETYLMPWIGVTSVLLLFIVYAYMYILWKAHSHAVRMIOGTOKSI	317
Db	175	--CPRECELFPLIPNDYLLSWLLFIAFLFSGIITYTGHVLMKAHQHVA-----SL	223
Qy	318	IIHTSEDKGVQVTRPDQARMA--IRLAKTLVLVLVLLIICWGPLLAIMVVDVFGKMKLI	375
Db	224	SGH--QDRQV----PGMARMLDVRLAKTLGLVLAVLLICWFPVLALMAHSLATTLSDQV	277
Qy	376	KTVFACSMCLLNSTVNPPIYALRSKDLR	405
Db	278	KKAFACSMCLCLNSMWNVPVIALRSGEIR	307

Search completed: January 6, 2006, 21:16:59  
Job time : 173 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 21:05:02 ; Search time 40 seconds  
(without alignments)

1135.358 Million cell updates/sec

Title: US-10-521-420-1

Perfect score: 2424

Sequence: 1 MKSILDGLADTTFTTTDL.....TVKIATKTVMSVSTDSAEAL 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Pir1.\*
- 2: Pir2.\*
- 3: Pir3.\*
- 4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2412	99.5	472	2 S17595	cannabinoid recept
2	2343.5	96.7	473	2 A33117	cannabinoid recept
3	734.5	30.3	347	2 S70364	cannabinoid recept
4	706	29.1	360	2 S36750	cannabinoid recept
5	363	15.0	364	2 JCS293	lysophosphatidic a
6	353	14.6	381	2 A35300	G protein-coupled
7	348.5	14.4	362	2 JC7559	spingosine 1-phos
8	346.5	14.3	378	2 JC5245	G protein-coupled
9	340	14.0	383	2 IS3870	Bdg-1 orphan recep
10	306.5	12.6	477	2 S71323	alpha-1A adrenergi
11	305	12.6	325	2 JC2193	melanocortin-5 rec
12	304	12.5	372	2 I49008	melanocortin-5 rec
13	296.5	12.2	352	2 JC1465	G protein-coupled
14	295.5	12.2	330	2 S40454	G protein-coupled
15	292	12.0	325	2 I46416	melanocyte-stimula
16	290.5	12.0	466	2 A35375	alpha-1C-adrenergi
17	289	11.9	330	2 A35689	G protein-coupled
18	288.5	11.9	363	2 S48697	probable G protein
19	285.5	11.8	297	2 C43265	adrenocorticotropi
20	285	11.8	466	2 IS7959	alpha-1C adrenergi
21	283	11.7	429	2 S65656	alpha-1C-adrenergi
22	283	11.7	466	2 JN0765	alpha-1C-adrenergi
23	283	11.7	499	2 S56567	alpha-1C-adrenergi
24	281	11.6	360	2 B46647	melanocortin recep
25	280	11.6	362	2 I65990	G protein-coupled
26	280	11.6	515	2 A40491	alpha-1-adrenergi
27	277	11.4	517	2 A45121	alpha-1B adrenergi
28	276	11.4	314	2 S71420	melanocortin 1 rec
29	276	11.4	314	2 S70005	melanocortin 1 rec

ALIGNMENTS

RESULT 1

S17595

cannabinoid receptor CB1 - human

C;Species: Homo sapiens (man)

C;Date: 03-Feb-1994 #sequence revision 03-Feb-1994 #text\_change 09-Jul-2004

C;Accession: S17595; S13668; A55879

R;Gerard, C.M.; Mollereau, C.; Vassart, G.; Parmentier, M.

Biochem. J. 279, 129-134, 1991

A;Title: Molecular cloning of a human cannabinoid receptor which is also expressed in

A;Reference number: S17595; MUID:92028798; PMID:1718258

A;Accession: S17595

A;Molecule type: mRNA

A;Residues: 1-472 <GER1>

A;Cross-references: UNIPROT:P21554; UNIPARC:UPI00000008AA; EMBL:X54937; NID:G25914; PFI

R;Gerard, C.; Mollereau, C.; Vassart, G.; Parmentier, M.

Nucleic Acids Res. 18, 7142, 1990

A;Title: Nucleotide sequence of a human cannabinoid receptor cDNA.

A;Reference number: S13668; MUID:91088303; PMID:2263478

A;Accession: S13668

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-472 <GER2>

A;Cross-references: UNIPARC:UPI00000008AA; EMBL:X54937; NID:G25914; PFI

R;Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.;

J. Biol. Chem. 270, 3726-3731, 1995

A;Title: An amino-terminal variant of the central cannabinoid receptor resulting from a

A;Reference number: A55879; MUID:95181329; PMID:7876112

A;Accession: A55879

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-124 <SHI>

A;Cross-references: UNIPARC:UPI000017C0AC; GB:X81120

C;Genetics:

A;Gene: GDB:CNRL1; CNR

A;Cross-references: GDB:127354; OMIM:114610

A;Map position: 6q14-q15

C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F;117-142/Domain: transmembrane #status predicted <TM1>

F;155-175/Domain: transmembrane #status predicted <TM2>

F;188-212/Domain: transmembrane #status predicted <TM3>

F;233-256/Domain: transmembrane #status predicted <TM4>

F;275-299/Domain: transmembrane #status predicted <TM5>

F;345-365/Domain: transmembrane #status predicted <TM6>

F;378-399/Domain: transmembrane #status predicted <TM7>

F;77,83/Binding site: carbohydrate (Asn) #status predicted

Query Match 99.5%; Score 2412; DB 2; Length 472;

Best Local Similarity 99.6%; Pred. No. 7,9e-192;

Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 MKSILDGLADTTFTTTDLTVGSGNDIQYEDIKGDMASKLGYPPQKPLTSFRGSPFOE 60

Db 1 MKSILDLADTTTTRTTTDLVYGSNDIQYEDIKGDMSKLGYPFPQKPLTSPRGSPFQ 60

QY 61 KMTAGNPOLVADQVNIITFYNKSLSSPKENEENIQCGENFNDIECFWVLPNSQQALAI 120

Db 61 KMTAGNPOLVADQVNIITFYNKSLSSPKENEENIQCGENFNDIECFWVLPNSQQALAI 120

QY 121 VLSLTGTFVLENLLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDFHFV 180

Db 121 VLSLTGTFVLENLLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDFHFV 180

QY 181 HRKDSRNVELFKLGGVTASTASVGLFLTAIRYISIHRLPLAYKRIIVTRPKAVAFCLM 240

Db 181 HRKDSRNVELFKLGGVTASTASVGLFLTAIRYISIHRLPLAYKRIIVTRPKAVAFCLM 240

QY 241 WTIAIVIAVLPGLWNCCKLOSCVDIIFPHIDTYLWFGVTSVLLLFIVVAYMYILWK 300

Db 241 WTIAIVIAVLPGLWNCCKLOSCVDIIFPHIDTYLWFGVTSVLLLFIVVAYMYILWK 300

QY 301 AHSHAVRMIQRGTQKSIITHTSDEGKVQVTRPDQARMDIRLAKTLVLVLIICWGPLL 360

Db 301 AHSHAVRMIQRGTQKSIITHTSDEGKVQVTRPDQARMDIRLAKTLVLVLIICWGPLL 360

QY 361 AIMVYDVGKWNKLIKTVFAFCSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCGTAQ 420

Db 361 AIMVYDVGKWNKLIKTVFAFCSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCGTAQ 420

QY 421 PLDNSMGDSDCILHKHANNAAVHRAAESCIKSTVKIAKVTMSVTDTSAEAL 472

Db 421 PLDNSMGDSDCILHKHANNAAVHRAAESCIKSTVKIAKVTMSVTDTSAEAL 472

RESULT 2

A33117

cannabinoid receptor CB1 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004

C;Accession: A33117; C55879

R;Matuoda, L.A.; Lolait, S.J.; Brownstein, M.J.; Young, A.C.; Bonner, T.I.

Nature 346, 561-564, 1990

A;Title: Structure of a cannabinoid receptor and functional expression of the cloned cDN

A;Reference number: A33117; MUID:90332039; PMID:2165569

A;Accession: A33117

A;Molecule type: mRNA

A;Residues: 1-473 <SHI>

A;Cross-references: UNIPROT:P20272; UNIPARC:UPI000012710C; GB:X55812; NID:g1552375; PIDN

R;Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.; Ca

J. Biol. Chem. 270, 3726-3731, 1995

A;Title: An amino-terminal variant of the central cannabinoid receptor resulting from al

A;Reference number: A55879; MUID:95181329; PMID:7876112

A;Accession: C55879

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-107 <SHI>

A;Cross-references: UNIPARC:UPI000017C808

C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 96.7%; Score 2343.5; DB 2; Length 473;

Best Local Similarity 96.8%; Pred. No. 3.7e-186; Indels 1; Gaps 1;

Matches 458; Conservative 5; Mismatches 9;

QY 1 MKSILDLADTTTTRTTTDLVYGSNDIQYEDIKGDMSKLGYPFPQKPLTSPRGSPFQ 60

Db 1 MKSILDLADTTTTRTTTDLVYGSNDIQYEDIKGDMSKLGYPFPQKPLTSPRGSPFQ 60

QY 61 KMTAGNPOLVPA-DQVNIITFYNKSLSSPKENEENIQCGENFNDIECFWVLPNSQQALAI 119

Db 61 KMTAGNSPLVPAAGDTTNIITFYNKSLSSPKENEENIQCGENFNDIECFWVLPNSQQALAI 120

QY 120 AVLSLTGTFVLENLLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDFHFV 179

Db 121 AVLSLTGTFVLENLLVLCVILHSRLCRPSYHFHIGSLAVADLLGSGVIFVYSFIDFHFV 180

QY 180 FHRKDSRNVELFKLGGVTASTASVGLFLTAIRYISIHRLPLAYKRIIVTRPKAVAFCL 239

Db 181 FHRKDSRNVELFKLGGVTASTASVGLFLTAIRYISIHRLPLAYKRIIVTRPKAVAFCL 240

QY 240 MTITAIIVIAVLPGLWNCCKLOSCVDIIFPHIDTYLWFGVTSVLLLFIVVAYMYILW 299

Db 241 MTITAIIVIAVLPGLWNCCKLOSCVDIIFPHIDTYLWFGVTSVLLLFIVVAYMYILW 300

QY 300 KAHSHAVRMIQRGTQKSIITHTSDEGKVQVTRPDQARMAIRLAKTLVLVLIICWGPLL 359

Db 301 KAHSHAVRMIQRGTQKSIITHTSDEGKVQVTRPDQARMDIRLAKTLVLVLIICWGPLL 360

QY 360 LAIMVYDVGKWNKLIKTVFAFCSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCGTA 419

Db 361 LAIMVYDVGKWNKLIKTVFAFCSMCLLNSTVNPPIIYALRSKDLRHAFRSMPPSCGTA 420

QY 420 QPLDNSMGDSDCILHKHANNAAVHRAAESCIKSTVKIAKVTMSVTDTSAEAL 472

Db 421 QPLDNSMGDSDCILHKHANNAAVHRAAESCIKSTVKIAKVTMSVTDTSAEAL 473

RESULT 3

S70364

cannabinoid receptor CB2, peripheral - mouse

N;Alternate names: G protein-coupled receptor CB2

C;Species: Mus musculus (house mouse)

C;Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004

C;Accession: S70364; S54163

R;Shire, D.; Calandra, B.; Rinaldi-Carmona, M.; Oustric, D.; Pessague, B.; Bonnin-Caban

Biochim. Biophys. Acta 1307, 132-136, 1996

A;Title: Molecular cloning, expression and function of the murine CB2 peripheral cannab

A;Reference number: S70364; MUID:96283804; PMID:8679694

A;Accession: S70364

A;Molecule type: mRNA

A;Residues: 1-347 <SHI>

A;Cross-references: UNIPROT:P47936; UNIPARC:UPI00000019CE; EMBL:X86405; NID:g791081; PI

C;Superfamily: melanocortin receptor

C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

Query Match 30.3%; Score 734.5; DB 2; Length 347;

Best Local Similarity 48.2%; Pred. No. 3.8e-53;

Matches 150; Conservative 53; Mismatches 89; Indels 19; Gaps 4;

QY 99 GENFMDIECFWVLPNSQQALAIIVLSLTGTFVLENLLVLCVILHSRLCRPSYHFHIGS 158

Db 16 GLEFNPMEKMYILSSQQIAVAVLCTLMGLLSALENMAVLYILSRRLRRKPSYLFISS 75

QY 159 LAVADLLGSGVIFVYSFIDFHFHVRKDSRNVELFKLGGVTASTASVGLFLTAIRYISI 218

Db 76 LAGADFLASVIFACNFVIFHVFHGVDSNAIFLLKIGSVTMTFTASVGSLLLTAVDRLCL 135

QY 219 HRPLAYKRIIVTRPKAVAFCLMWTITAIIVIAVLPGLWNCCKLOSCVDIIFPHIDTYLWF 278

Db 136 CYPPTYKALVTRGRALVLCVMVLSALISYPLMGWTC--CPSPCSELPFLIPNDYLLG 193

QY 279 WIGVTSVLLLFIVVAYMYILWKAHSHAVRMIQRGTQKSIITHTSDEGKVQVTRPDQARMA 338

Db 194 WLLFTAILFSGIITYTYGVVLWKAHRHVATLAEH-----QDRQV----PGIARMR 238

QY 339 --IRLAKTLVLVLIICWGPLLAIWYDVGKWNKLIKTVFAFCSMCLLNSTVNPPII 396

Db 239 LDVRLAKTLGLVLVLIICWFFPALALMGHSLVTLTSDQVKEAFAPFCMLCLVNSMVPPII 298

QY 397 YALRSKDLRHA 407

Db 299 YALRSGEIRSA 309

RESULT 4

S36750

cannabinoid receptor CB2 - human

N;Alternate names: cannabinoid receptor, peripheral

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1994 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-1999





Qy 409 -----BSMPFSCGTAQPLDNS 425  
Db 324 IRIISCKCPNGDSAGFKRPIIPGMBFSPRSKSDNS 359

RESULT 10  
S71323  
alpha-1A adrenergic receptor - Japanese medaka  
C:Species: Oryzias latipes (Japanese medaka)  
C>Date: 23-Jul-1997 #sequence\_revision 01-Aug-1997 #text\_change 17-Mar-1999  
C:Accession: S71323  
Eur J. Biochem. 235, 501-507, 1996  
R:Yasuoaka, A.; Abe, K.; Arai, S.; Emori, Y.  
A>Title: Molecular cloning and functional expression of the alpha-1-A-adrenoceptor of Me  
A:Reference number: S71323; MUID:96184522; PMID:8654394  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-477 <VAS>  
A:Cross-references: UNIPARC:UPI00001778AD; EMBL:D63859  
A:Note: it is uncertain whether Met-1 or Met-8 is the initiator  
C:Superfamily: vertebrate rhodopsin

Query Match 12.6%; Score 306.5; DB 2; Length 477;  
Best Local Similarity 27.4%; Pred. No. 1.5e-17;  
Matches 93; Conservative 69; Mismatches 130; Indels 47; Gaps 13;

Qy 107 CFVNLNPS-QQLAIAVLSLTIGTFT---VLENLLVLCVILHSRLCRPSYHFTGSLAVA 162  
Db 20 CSHVLAPELNTKAVVLGMVLGIFLFGVIGNILVILSVCHRHQ-LTVTYFVNLAVA 78

Qy 163 D-LLGSVTFVYSDFDFVHRKDRNVFLKGGV-TASFTASVGSFLTAIARYSIHR 220  
Db 79 DLLLSSTVLPSAI-FEILDRWFRGVFCNIWAAVDVLCCTASIMSLCVISVDRIYGVSY 137

Qy 221 PLAKRIVTRPKAVAPCLMTIAIVAPLPGNW---CEKLSQVCS-----DIPPHI 271  
Db 138 PLRPAIMTKRALLVAMLLVLSIIIGLFGKPEAPDETVCKITEPFGVAISAV 197

Qy 272 DETYLMFWIGVTSFLFVYAYMYILWKAHSHAVR---MIQRTGQKSIHH-----T 321  
Db 198 GSFYL-----PLAILAMCYRVVVAQESRGLKEQKIEKSDSEQVILRMHGRNTV 250

Qy 322 SEDGKVQVTRPDQARMAIRL-----AKTLVLVLVLCVGLLAIMYVDVFGKWN 372  
Db 251 SEDEALR----SRTHFALRLKLFGRKKAATLGIWVGCFLVCLWLPFLVLPFGSIFPAY 306

Qy 373 KLIKTVPFCSMLCLNSTVNPPIIYALRSKDLRHAFRSM 411  
Db 307 RPSDTVFKITFWLGVFNSCINPIIYLCNSQBFKAQFSL 345

RESULT 11  
JC2193  
melanocortin receptor, MC5 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1993 #sequence\_revision 20-Aug-1994 #text\_change 09-Jul-2004  
C:Accession: JC2193  
R:Griffon, N.; Mignon, V.; Facchinetti, P.; Diaz, J.; Schwartz, J.C.; Sokoloff, P.  
Biochem. Biophys. Res. Commun. 200, 1007-1014, 1994  
A>Title: Molecular cloning and characterization of the rat fifth melanocortin receptor.  
A:Reference number: JC2193; MUID:94234987; PMID:8179577  
A:Accession: JC2193  
A:Molecule type: DNA  
A:Residues: 1-325 <GRI>  
A:Cross-references: UNIPROT:P35345; UNIPARC:UPI000012ED53; GB:L27081; NID:G435606; PIDN:  
C:Superfamily: melanocortin receptor  
C:Keywords: G protein-coupled receptor; receptor; transmembrane protein  
F:37-63/Domain: transmembrane #status predicted <TM1>  
F:73-98/Domain: transmembrane #status predicted <TM2>  
F:117-138/Domain: transmembrane #status predicted <TM3>  
F:159-179/Domain: transmembrane #status predicted <TM4>

F:193-211/Domain: transmembrane #status predicted <TM5>  
F:240-263/Domain: transmembrane #status predicted <TM6>  
F:276-297/Domain: transmembrane #status predicted <TM7>

Query Match 12.6%; Score 305; DB 2; Length 325;  
Best Local Similarity 27.5%; Pred. No. 1.2e-17;  
Matches 95; Conservative 66; Mismatches 118; Indels 66; Gaps 13;

Qy 92 NEENIQGENFMDEICPMVNLNPS---QQLAIAV-LSLTIGTFTVLENLLVLCVILHSRL 147  
Db 17 SEDNI-LGQN-----VNNKSACEDMGIAVEFTLGLVLSLEENILVIGAVKKNL 67

Qy 148 RCRPSYHFIGSLAVADLLGSVIFYSPIDFVHRKD--SRNVFLKGGV-----TASF 200  
Db 68 H-SPMYFFVGSGLAVADMVLSMSNAWETITVILNNKHVIAADTFVRHIDNVFDSMICISV 126

Qy 201 TASVGSFLTAIARYSIHRPLAYKRIVTRPKAVAVACLMTIAIVAVLPLGWNCEKL 260  
Db 127 VASMSLLAIAVDRIYITFYALRYHHIMTARRSGVIIACIWTFCISGIVFIIYYE-SKY 185

Qy 261 QSVCSDFPHIDETVLMFWIGVTSVLLLFVYAYMYILWKAHSHAVRMIQRTGQKSIHH 320  
Db 186 VIVC-----LISMPFTMLFFWVSLYIHMFLARNVHVKRIASPRNSV--- 228

Qy 321 TSEDGKVQVTRPDQARMAIRLAKTLVLVLVLCVGLLAIMYVDVFGKWNKLIKTVFA 380  
Db 229 -----RQRASMKGAITLWMLGIFVCSPPFLHLILMISCPQN-----VYC 270

Qy 381 FCSM-----LCLNSTVNPPIIYALRSKDLRHAFRSMFPSCG 417  
Db 271 ACFSYFNMYLLILMCSVIDPLIYALRSQEMRTFKELI-CCHG 314

RESULT 12  
I49008  
melanocortin-5 receptor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I49008; A54245; JC2244  
R:Pathi, Z.; Iben, L.G.; Parker, E.M.  
Neurochem. Res. 20, 107-113, 1995  
A>Title: Cloning, expression, and tissue distribution of a fifth melanocortin receptor  
A:Reference number: I49008; MUID:95258173; PMID:7739752  
A:Accession: I49008  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-372 <RES>  
A:Cross-references: UNIPROT:P41149; UNIPARC:UPI000016384D; EMBL:U09354; NID:G522165; P:  
R:Labbe, O.; Desarnaud, P.; Eggerickx, D.; Vassart, G.; Parmentier, M.  
Biochemistry 33, 4543-4549, 1994  
A>Title: Molecular cloning of a mouse melanocortin 5 receptor gene widely expressed in  
A:Reference number: A54245; MUID:94213827; PMID:8161509  
A:Accession: A54245  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 48-120, 'P', 122-372 <LAB>  
A:Cross-references: UNIPARC:UPI0000027516; GB:X76295; NID:G498973; PIDN:CAA53943.1; P:  
A:Experimental source: Clone HGMPO1B  
A:Note: sequence extracted from NCBI backbone (NCBI:145988)  
R:Gantz, I.; Shimoto, Y.; Konda, Y.; Miwa, H.; Dickinson, C.J.; Yamada, T.  
Biochem. Biophys. Res. Commun. 200, 1214-1220, 1994  
A>Title: Molecular cloning, expression, and characterization of a fifth melanocortin r  
A:Reference number: JC2244; MUID:94241974; PMID:8185570  
A:Accession: JC2244  
A:Molecule type: DNA  
A:Residues: 48-372 <GAN>  
A:Cross-references: UNIPARC:UPI0000147322; GB:L22527; NID:G468377; PIDN:AAA21337.1; P:  
C:Comment: This protein responds to melanocortins with an increase in intracellular c  
C:Superfamily: melanocortin receptor  
C:Keywords: receptor; transmembrane protein  
F:84-109/Domain: transmembrane #status predicted <TM1>  
F:121-144/Domain: transmembrane #status predicted <TM2>  
F:164-185/Domain: transmembrane #status predicted <TM3>

F;206-226/Domain: transmembrane #status predicted <TM4>  
F;232-257/Domain: transmembrane #status predicted <TM5>  
F;287-310/Domain: transmembrane #status predicted <TM6>  
F;324-344/Domain: transmembrane #status predicted <TM7>

Query Match 12.5%; Score 304; DB 2; Length 372;  
Best Local Similarity 26.4%; Pred. No. 1.8e-17;  
Matches 84; Conservative 67; Mismatches 123; Indels 44; Gaps 9;

QY 115 QQLAIYVLSLTGTFVLENLVLCVILHSRLCRPSYHFGSLAVADLLGSVIFVYS 173  
DB 81 EEMGIAYEVFTLGLVSLLENLIVGAIYKKNLH-SPMYFYVGSVAVADMLVMSNAWE 139  
QY 174 FIDEVHFHKD--SRNVFLKGGV-----TASFTASVGLFTATARIYISHRPLAYKR 226  
DB 140 TTVIYLNNKHLVIADTFVRHIDNVFDSMICISVAVSMCSLLAIAVDRIITFYALRYHH 199  
QY 227 IVTRPKAVVAFCLMWTIAIVLPLLGWNCCKLOQSVCSDFPHIDETVLMFWIGVTSVL 286  
DB 200 IMTARSGVLIACIWTFCISGVFIYYE-SKYVVIC-----LISMFFTM 244  
QY 287 LFIIVYAVYMLWKASHAVRMQRTQKSIHTSEDGKQVQVTRPDQARMAIRLAKTLV 346  
DB 245 LFFWVSLYTHMFLAARNVHKRIAASPRYSVRQRTSMKGA-----TLT 288  
QY 347 LILVLIICWGPLLAIMVYDFGKMKLIKTVFAFCSM---LCLLNSTVNPPIIYALRSKD 403  
DB 289 MLLGIFIVCSPPFLHILMISCPQNYCSPMSYFNWVLLIMCNSVIDPLIYALRSQE 348  
QY 404 LRHAFRSMPPSCBGTAPQ 421  
DB 349 MRRTFKEIV-CHGFRPP 365

## RESULT 13

JC1465  
Probable G protein-coupled receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C:Accession: JC1465  
R;Okazaki, H.; Ishizaka, N.; Sakurai, T.; Kurokawa, K.; Goto, K.; Kumada, M.; Takuwa, Y.  
Biochem. Biophys. Res. Commun. 190, 1104-1109, 1993  
A:Title: Molecular cloning of a novel putative G protein-coupled receptor expressed in b  
A:Reference number: JC1465; MUID:93176155; PMID:8382486  
A:Accession: JC1465  
A:Molecule type: mRNA  
A:Residues: 1-352 <OKA>  
A:Cross-references: UNIPROT:P47752; UNIPARC:UPI000000007D; GB:AB016931; NID:g3445557; PT  
A:Experimental source: aortic smooth muscle  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot  
F;35-59/Domain: transmembrane #status predicted <TM1>  
F;67-95/Domain: transmembrane #status predicted <TM2>  
F;110-128/Domain: transmembrane #status predicted <TM3>  
F;148-173/Domain: transmembrane #status predicted <TM4>  
F;190-210/Domain: transmembrane #status predicted <TM5>  
F;224-255/Domain: transmembrane #status predicted <TM6>  
F;272-293/Domain: transmembrane #status predicted <TM7>  
F;19/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;142,145,218,219,329,330,331,332/Binding site: phosphate (Ser) (covalent) #status predi  
F;313/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 12.2%; Score 296.5; DB 2; Length 352;  
Best Local Similarity 29.4%; Pred. No. 6.9e-17;  
Matches 99; Conservative 48; Mismatches 117; Indels 73; Gaps 12;

QY 92 NEENIQGFENFMDIECFMVLNPSQOLAIVLSLTGTFVLENLVLCVILHSRLCRP 151  
DB 10 NPEKVQEHYNYTKETLDMQETPSRKVASAFI-IILCCAIWENLVLLIAVARNSKEFH-SA 67  
QY 152 SYHPIGSLAVADLLGSVIFV-----YSFIDEVHFHKDSRNVFLFKGGVTSAPT 201  
DB 68 MYLFLGNLAASDLLAGVAFVNTLSPVTLSTPLQWFARE-----GSAFITLIS 117

QY 202 ASVGSGLFTATARIYISHRPLAYKRIVTRPKAVVAFCLM---WTIAIVIAVLPLLGWNCE 258  
DB 118 ASVFSLLAIAIERQVAI---AKVKLYGSDKSCRMMLIGASWLSLILGGLPILGWNC 173  
QY 259 KLOQSVCSDFPHIDETVLMFWIGVTSVLLLFIVYAVYMLWKASHAVRMQRTQKSI 317  
DB 174 DHLEACSTVLPYAKHYVLCVVTFISVILLAIYALYVRIYFVVRSSHA----- 221  
QY 318 IIHTSEDGKQVQVTRPDQARMAIRLAKTLVLIIVLVLICWGPLLAIMVYD---VF 368  
DB 222 -----DVAGP-----QTLALAKTVIIVLGVTIICWLPAPFSLILLDSTCFVRACPVL 267  
QY 369 GRMNKLIKTVFAFCSMCLLNSTVNPPIIYALRSKDLR 405  
DB 268 YKAHYF-----FAFAT-----LNSLLANPVIYTWESRDLR 296

## RESULT 14

S40454  
G protein-coupled receptor GPCR21 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 19-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S40454  
R;Saeiki, Y.; Ueno, S.; Mizuno, R.; Nishimura, T.; Fujimura, H.; Nagai, Y.; Yanagihara, H.  
FEBS Lett. 336, 317-322, 1993  
A:Title: Molecular cloning of a novel putative G protein-coupled receptor (GPCR21) whic  
A:Reference number: S40454; MUID:94085630; PMID:8262253  
A:Accession: S40454  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-330 <SAE>  
A:Cross-references: UNIPROT:P35413; UNIPARC:UPI0000025C17; GB:D21062; NID:g455487; PIDN  
C:Superfamily: melanocortin receptor  
C:Keywords: G protein-coupled receptor; phosphoprotein; transmembrane protein

Query Match 12.2%; Score 295.5; DB 2; Length 330;  
Best Local Similarity 28.0%; Pred. No. 7.8e-17;  
Matches 90; Conservative 61; Mismatches 125; Indels 45; Gaps 13;

QY 111 LNPQOLAIVLSLTGTFVLENLVLCVILHSRLCRPSYHFGSLAVADLLGSVIF 170  
DB 36 LLPSPRAMDVLCIS-GTLVSCENALVVAIVGTFAFRA-PMFLVGLSVAVADLLAGLGL 93  
QY 171 VYYSF-IDFHVHFHKDSRNVFLFKGGVTSFTASVGSGLFTATARIYISHRPLAYKRIVT 229  
DB 94 VLHFAADFCL-----GSPENSLMLVGLVLAWAFATASIGSLAITVDRLSYLNALTYVSETT 149  
QY 230 RPKAVVAFCLMWTIAIVIAVLPLLGWNCCKLOQSVCSDFPHIDETVLMFWIGVTSVLLLF 289  
DB 150 VTRTTYMLALVWVGALGLVPLAWNCRDGLTTCGVVYP-LSKNHL-----VVLAIAPF 203  
QY 290 IYVAYMYLWKASHAVRMQRTQK-SIIHTSEDGKQVQVTRPDQARMAIRLAKTLVLI 348  
DB 204 MVFG---IMLQYACICRVCRHAQQIALQORLLPASHVATRKGFIA-----TLAVV 252  
QY 349 LVVLIICWGPLLAIMVYDFGKMKLIKTVFAFCSM-LCLLNSTVNPPIIYALRSKDLRHA 407  
DB 253 LGAPACW---LPFTVYCLLGADS---PRLYTYTLTLPATYNSMINPVIYAFENQDVQKV 307  
QY 408 -----FRSMFPPS 414  
DB 308 LWAICCCSTSKIPFRSRSPS 328

## RESULT 15

I46416  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C>Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: I46416; S43120  
R;Barrett, P.; MacDonald, A.; Helliwell, R.; Davidson, G.; Morgan, P.  
J. Mol. Endocrinol. 12, 203-213, 1994



A:Title: Cloning and expression of a new member of the melanocyte-stimulating hormone re  
A:Reference number: I46416; MUID:94338523; PMID:8060485  
A:Accession: I46416  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-325 <BAR>  
A:Cross-references: UNIPROT:P41983; UNIPARC:UPI000012ED54; EMBL:Z31369; NID:9467692; PII  
A>Note: submitted to the EMBL Data Library, March 1994  
C:Superfamily: melanocortin receptor

Query Match		12.0%; Score 292; DB 2; Length 325;
Best Local Similarity		28.4%; Pred. No. 1.5e-16;
Matches		99; Conservative 61; Mismatches 88; Indels 100; Gaps 17;
Qy	112	NPQQIAIAY-LSLTGTFVLENLVLCVILHSRLCRPSYHFGISLAVADLLGSV-- 168
Db	31	SPCEDMGIAVEVFLALGLISLLENILVIGAVRNRLHI-PMYFFVGSLAVADMLVSLN 89
Qy	169	----IFVYSFIDFHVHRKDS---RNVFLKLGGVTFASVGSFLTAIARYISHR 220
Db	90	FWETIYLLTNKHLVWADASVRHLDNVP---DSMICISVVASMCSLAIAVDVRYTIFC 146
Qy	221	PLAYKRIVT--RPKAVV---AFCLMWTIAIVAVLPLGWNCKEKLQSVCSDFP-PHIDE 273
Db	147	RLAYQRIHTGRSGAIIAGIWAFC-----TSCGTVFIVVYES 183
Qy	274	TYL-----MFWIGVTSVLLFIVYAYWYILWKAHSHAVRMIOGTOKSIIHTSEDKV 327
Db	184	TYVVVCLIAMF--LTMLLMASLYTHMFLARTH---VRR-----AALPGHSSV--- 228
Qy	328	QVTRPDQARMAIRLAKTLVLILVLLICWGPLLAIMVVDVFGKMKLIKTVFAFCS--- 383
Db	229	-----KQRTGVKGAITLMLLGVFIICWAPFFLHLIMISCPQN-----LYCSCFMS 275
Qy	384	-----MLCLLNSTVNPPIIYALRSKOLRHAFRSM-----FPS 414
Db	276	HNNYLLIMCNSVIDPLIYAFRSQEMKTFKEIVCFQGFRTPCRPFS 323

Search completed: January 6, 2006, 21:14:03  
Job time : 42 secs

**This Page Blank (uspto)**